

Random Fields Simplified: Preserving Marginal Distributions, Correlations, and Intermittency, with applications from rainfall to humidity

Simon Michael Papalexiou^{1,2,3} and Francesco Serinaldi^{4,5}

¹Department of Civil, Geological and Environmental Engineering, University of Saskatchewan, Canada; ²Global Institute for Water Security, Saskatchewan, Canada; ³Faculty of Environmental Sciences, Czech University of Life Sciences Prague ⁴School of Engineering, Newcastle University, Newcastle Upon Tyne NE1 7RU, UK; ⁵Willis Research Network, 51 Lime St., London, EC3M 7DQ, UK.

Abstract Nature manifests itself in space and time. The spatiotemporal complexity of processes such as precipitation, temperature, wind, etc., does not allow purely deterministic modelling. Spatiotemporal random fields have a long history in modelling such processes and yet a single unified framework offering the flexibility to simulate processes that may differ profoundly does not exist. Here we introduce a blueprint to efficiently simulate spatiotemporal random fields that preserve any marginal distribution, any valid spatiotemporal correlation structure, as well intermittency. We suggest a set of parsimonious yet flexible marginal distributions and provide a rule-of-thumb for their selection. We propose a new and unified approach to construct flexible spatiotemporal correlation structures by combining copulas and survival functions. The versatility of our framework is demonstrated by simulating conceptual cases of intermittent precipitation, double-bounded relative humidity, and temperature maxima fields. As a real-world case we simulate daily precipitation fields. In all cases, we reproduce the desired properties. In an era characterized by advances in remote sensing and increasing availability of spatiotemporal data, we deem this unified approach offers a valuable and easy-to-apply tool for modelling complex spatiotemporal processes.

Key Points

- Versatile and easy modelling of random spatiotemporal fields
- The generated fields reproduce any marginal distribution, correlation structure, and intermittency
- New spatiotemporal correlation structures based on copulas and survival functions

1. Introduction: Brief Literature Review on Random Fields

Hydro-climatic, environmental and geophysical processes such as precipitation, temperature, relative humidity and surface wind fields are characterized by spatial and temporal variability (Sun et al., 2012). The complexity of the dynamics of such processes evolving in both space and time often prevents the use of deterministic models based on the governing equations. Therefore, stochastic models represent viable alternatives that are generally parsimonious in terms of the number of parameters and correspond to processes mimicking many patterns of the observed spatiotemporal variability, without necessarily following the underlying governing equations (Kyriakidis & Journel, 1999).

Among stochastic models, random fields provide the most appropriate descriptions of natural processes evolving in space and time. Let Ω be the sample space of a random experiment, \mathbb{R}^n the n -dimensional (spatial) Euclidean space, and $T \subseteq \mathbb{R}_{+,0}^1$ the 1-dimensional (temporal) Euclidean space; a spatiotemporal random field is a stochastic process, $X(u; \mathbf{s}, t)$, where $u \in \Omega$, $(\mathbf{s}, t) \in \mathbb{R}^n \times T$, i.e. a collection of random variables $\{X_1, \dots, X_n\}$ at points (coordinates) $\mathbf{s}_1, \dots, \mathbf{s}_n, t$, over the Cartesian product $\mathbb{R}^n \times T$ (see Christakos, 2005 Chapter 2.2 and 5.3 for formal definitions and more details).

The study of spatial and spatiotemporal random field models has quite a long history (Sun et al., 2012; Vanmarcke, 1983). Nonetheless, it is still a very active area of research because of the new challenges related to analysis and modelling of large datasets coming from remote sensors (Sun et

al., 2012), or the study of spatial extreme values (Bacro & Gaetan, 2012) in a variety of applied fields, such as epidemiology and health risk assessment (Christakos & Hristopulos, 1998; Angulo et al., 2013; Fei et al., 2019; He et al., 2019), meteorology, environmental monitoring, and design (e.g. Genton, 2007; Jona-Lasinio et al., 2012; Kleiber, 2016; Li et al., 2008; Nevat et al., 2015; Sampson & Guttorp, 1992; Youngman & Stephenson, 2016), geotechnics and reservoir engineering (Hardy & Beier, 1994; Ronold Knut O., 1990), image analysis, computer vision and medical imaging (Winkler, 1995; Karimaghloo et al., 2016, and references therein), finance and insurance (e.g. Goldstein, 2000; Lescourret & Robert, 2006; Evstigneev & Zhitlukhin, 2013), and many others. Since spatiotemporal modelling deals mainly with the second order properties of random fields evolving temporally over a given spatial domain, covariance functions describing the interactions between spatial and temporal components are fundamental for estimation, prediction and simulation (Alegria & Porcu, 2017; Gneiting et al., 2006). Constructing valid positive definite spatiotemporal covariance functions is sometimes closely related to simulation methods (Schlather, 2012).

First and second order properties (i.e. mean, variance and covariance) are sufficient statistics when the process is Gaussian. However, Gaussian spatial or spatiotemporal models rarely fit the observations of environmental processes. Therefore, non-Gaussian options should be considered. The approaches proposed for building non-Gaussian spatial processes can be grouped in two main classes (Allard, 2012): (1) methods involving the transformation of the data into values that are compatible with a Gaussian process or another tractable process (Chilès & Delfiner, 2009; Wackernagel, 2003); and (2) methods assuming that the random field follows specific processes, such as χ^2 or t -Student (e.g. Adler & Taylor, 2007), Gamma (Wolpert & Ickstadt, 1998) or skew-normal (Allard, 2012). However, these approaches are not flexible and general enough to model complex processes such as precipitation or wind. These processes are intermittent, and their conditional distribution are characterized by large variability in shape.

Although a variety of simulation algorithms are available in the literature (see Christakos, 2005; Lantuejoul, 2013; Schlather, 2012 for a review), most of them are devised for random processes with Gaussian marginal or joint distribution and enable only simulation of spatially correlated but temporally independent random fields (Schlather, 1999). Therefore, independently of the strategy used to build non-Gaussian random fields, their simulation often relies on the simulation of parent Gaussian fields. This method, which is known in spatiotemporal stochastics literature as the nonlinear transformation approach (Christakos, 2005 pp. 332-334; Liou et al., 2011), comprises the transformation of the marginal distribution via the so-called normal quantile/score transform (also known as Gaussian anamorphosis, and marginal-back transform), and the transformation (inflation) of the covariance function. The latter aims to obtain the covariance of the parent Gaussian process yielding the desired covariance of the target process after back applying the marginal-back transformation. However, we stress again these schemes lack the generality to simulate intermittent processes, or processes with mixed-type marginals with probability mass at arbitrary points, and reproduce any spatiotemporal correlation structure. Yet their most prominent disadvantage is the use of cumbersome numerical approaches to estimate the inflated Gaussian correlations that typically are constrained in preserving the lag-1 or lag-2 correlations.

Following the rationale of nonlinear transformation approach and building on the methodological results reported by Papalexiou (2018), in this study we introduce a framework to efficiently simulate spatiotemporal random fields with prescribed marginal distribution and spatiotemporal correlation suitable to describe a variety of environmental variables evolving over spatiotemporal lattices.

The paper is structured as follows. Section 2 presents the key elements of the modelling and simulation methods consisting of a set of marginal distributions, nonlinear transformation, and

spatiotemporal correlation functions. We suggest a set of marginal distributions that are based on entropy maximization and simple rules which are flexible enough to model a great variety of natural and non-natural processes. The nonlinear transformation method to obtain the Gaussian spatiotemporal correlation structure is optimized by an efficient fitting and parametrization procedure relying on a simple correlation transformation function. We also suggest an alternative approach to build potentially permissible (valid) spatiotemporal correlation functions based on copulas and univariate survival functions, discussing the case of separable (independent) spatial and temporal correlation functions as well. In Section 3, we report some representative examples demonstrating the versatility of the proposed approach and its capability to simulate virtual fields of precipitation with discrete-continuous marginal distributions, relative humidity with double bounded beta marginal distributions, and daily maximum temperatures. As a real-world example, we show in Section 4 the simulation of spatiotemporal rainfall fields reproducing rainfall records over a spatial lattice of 0.05° resolution extracted from CHIRPS (Climate Hazards Group InfraRed Precipitation with Station data). An extensive discussion in Section 5 is followed by perspectives and conclusions in Section 6.

2. Random Fields Preserving Marginal and Correlations

2.1 Probability distributions for fields

In the most general case, and in layman's terms, a spatiotemporal field is a collection of a finite (or infinite) number of random variables (r.v.'s) interrelated with each other. The properties of these r.v.'s may vary in time and space. Thus, their marginal distributions are an essential component of the field, with the other being dependence structures among them. We could say that the marginal distribution embodies the "potential" of a process—it dictates the frequency or the likelihood of its different magnitudes (values) to occur; it governs the frequency and the magnitude of extremes events. Thus, choosing a marginal distribution is crucial, and yet the selection is often based on performance criteria that may lead to choices neglecting even the variable's feasible range. Random variables of natural processes such as precipitation, wind, discharge, etc., take positive values and are typically positively skewed, or heavy tailed (at fine spatiotemporal scales); however, it is not uncommon in the practice to use distributions with inconsistent support, or inconsistent tails. For example, distributions having a location parameter (except those defined in the whole real line) end up having a lower bound or even an upper bound. Widely used distributions such as the Gamma and the Exponential have thin tails and cannot model adequately extremes of heavy-tailed variables; for example, global analyses show that nonzero daily precipitation has, in general, heavier tails than the exponential (see e.g., Papalexiou et al., 2013; Serinaldi & Kilsby, 2014).

The number of distributions that can be formed, at least theoretically, is infinite. The common practice of fitting many distributions (now facilitated greatly by using software) and selecting the best fitting option, as aforementioned, may lead to inconsistencies. In this respect, Papalexiou & Koutsoyiannis (2012) derived consistent distributions for positive geophysical and hydroclimatic r.v.'s based on theoretical reasoning and using entropy measures. Yet as a general rule of thumb, we propose that the continuous distributions to be used in practice should be: (1) consistent with the valid range of the variable under study, (2) flexible enough to model heavy tails, and, J- and Bell-shaped probability densities, and (3) have the least number of parameters necessary to fulfil the previous two requirements.

Based on the previous three-point mandate, we suggest the Generalized Gamma (\mathcal{GG}) (Stacy, 1962; Papalexiou & Koutsoyiannis, 2012), and the Burr type III and XII distributions ($\mathcal{B}\mathcal{r}\mathcal{XII}$) (Burr, 1942; Tadikamalla, 1980; Papalexiou, 2018) as general and flexible candidate distributions. We also propose here a generalization of the Beta of the second kind (\mathcal{BII}) which for $\gamma_2 \rightarrow 0$ becomes the

celebrated Gamma distribution (\mathcal{G}). The probability densities of the \mathcal{GG} and the \mathcal{BII} distributions, and the distribution functions of the \mathcal{BrXII} and \mathcal{BrXIII} are given, respectively, by

$$f_{\mathcal{GG}}(x; \beta, \gamma_1, \gamma_2) = \frac{\gamma_2}{\beta \Gamma(\gamma_1/\gamma_2)} \left(\frac{x}{\beta}\right)^{\gamma_1-1} \exp\left(-\left(\frac{x}{\beta}\right)^{\gamma_2}\right) \quad (1)$$

$$f_{\mathcal{BII}}(x; \beta, \gamma_1, \gamma_2) = \frac{\gamma_2^{\gamma_1}}{\beta \mathcal{B}(\gamma_1, 1/\gamma_2)} \left(\frac{x}{\beta}\right)^{\gamma_1-1} \left(1 + \gamma_2 \frac{x}{\beta}\right)^{-\gamma_1 - \frac{1}{\gamma_2}} \quad (2)$$

$$F_{\mathcal{BrXII}}(x; \beta, \gamma_1, \gamma_2) = 1 - \left(1 + \gamma_2 \left(\frac{x}{\beta}\right)^{\gamma_1}\right)^{-\frac{1}{\gamma_1 \gamma_2}} \quad (3)$$

$$F_{\mathcal{BrXIII}}(x; \beta, \gamma_1, \gamma_2) = \left(1 + \frac{1}{\gamma_1} \left(\frac{x}{\beta}\right)^{-\frac{1}{\gamma_2}}\right)^{-\gamma_1 \gamma_2} \quad (4)$$

These distributions, defined in $(0, \infty)$, have a scale parameter $\beta > 0$, and two shape parameters $\gamma_1 > 0$ and $\gamma_2 > 0$ that control, respectively, the left and right tail. The product moments of these distributions are given in Annex A and can be used to estimate the mean, variance, skewness, etc. If simpler two-parameter models, such as the Pareto II, Lognormal, Weibull, Gamma, etc., can describe the data and the extremes adequately, they should be preferred. Also, in cases dealing with extremes such as annual maxima or peaks over threshold the well-known extreme value distributions (Generalized Extreme Value distribution (\mathcal{GEV}) and Generalized Pareto (\mathcal{GP})) should be considered as they are supported by theoretical justification.

The marginal distributions used to describe hydroclimatic processes and fields can be continuous but also discrete, binary, or mixed type (for a brief review on probability models used in hydroclimate see Papalexiou, (2018)). The foregoing strictly continuous distributions for r.v.'s defined in $(0, \infty)$, can be modified to suit many fundamental hydroclimatic processes such as precipitation or wind, which are intermittent in space and time. In such cases the unconditional marginal distribution $F_X(x)$ has a probability mass at zero p_0 (or else it is zero inflated) with distribution, quantile, mean and standard deviation expressions given, respectively, by

$$F_X(x) = (1 - p_0)F_{X|X>0}(x) + p_0 \quad x \geq 0 \quad (5)$$

$$x_u = Q_X(u) = \begin{cases} 0 & 0 \leq u \leq p_0 \\ Q_{X|X>0}\left(\frac{u - p_0}{1 - p_0}\right) & p_0 < u \leq 1 \end{cases} \quad (6)$$

$$\mu_X = (1 - p_0)\mu_{X|X>0} \quad (7)$$

$$\sigma_X^2 = (1 - p_0)\sigma_{X|X>0}^2 + p_0(1 - p_0)\mu_{X|X>0}^2 \quad (8)$$

where $X|X > 0$ denotes the conditional r.v. (positive in this case).

The use of mixed-type distributions in calculating directly the parent-Gaussian correlations, introduced in Papalexiou (2018), is readily extended here to simulate intermittent spatiotemporal fields. This allows the simulation of intermittency in space, in time, or in both. We can also generalize this scheme for mixed-type marginals inflated at arbitrary points by modifying accordingly the expressions given in Equations (5) to (8) and using them in Equations (13) and (14). A useful case of mixed-type distributions inflated in two points regards processes where the random variable has a probability mass at its minimum and maximum values (it could also be arbitrary chosen low and

high thresholds). Particularly, if the minimum and maximum values, x_1 and x_2 , of a random variable X , have probability masses p_1 and p_2 , respectively, and $X|X \in (x_1, x_2)$ is continuous, then

$$F(x) = \begin{cases} (1 - p_1 - p_2)F_{X|X \in (x_1, x_2)}(x) + p_1 & x < x_2 \\ 1 & x \geq x_2 \end{cases} \quad (9)$$

$$x_u = Q_X(u) = \begin{cases} x_1 & 0 \leq u \leq p_1 \\ \left(\frac{u - p_1}{1 - p_1 - p_2} \right) & p_1 < u \leq 1 - p_2 \\ x_2 & 1 - p_2 < u \leq 1 \end{cases} \quad (10)$$

$$\mu_X = \sum_{i=1}^2 p_i x_i + \left(1 - \sum_{i=1}^2 p_i \right) \mu_{X|X \in (x_1, x_2)} \quad (11)$$

$$\sigma_X^2 = \sum_{i=1}^2 p_i x_i^2 + \left(1 - \sum_{i=1}^2 p_i \right) \sigma_{X|X \in (x_1, x_2)}^2 + \left(1 - \sum_{i=1}^2 p_i \right) \mu_{X|X \in (x_1, x_2)}^2 - \mu_X^2 \quad (12)$$

Thus, the unconditional distribution function $F(x)$ and the quantile function $Q_X(u)$ become piecewise functions while the mean and variance of X can be directly estimated. A real-world case could be the water depth of small lakes or ponds that dry up (probability mass at zero depth) or overspill for long periods due to inflows or long-lasting precipitation events (probability mass at maximum depth); the in-between states can be described by a continuous distribution with values ranging in (min, max).

2.2 Correlation Transformations of the parent-Gaussian Fields

Let $\{X(\mathbf{s}, t): \mathbf{s} \in \mathbb{R}^2, t \in [0, \infty)\}$ be a spatiotemporal target process having a prescribed marginal distribution $F_X(x)$ with finite variance, and a valid (positive definite) spatiotemporal correlation structure (STCS) $\rho_X(\delta, \tau)$; where δ is the Euclidean distance between the points \mathbf{s}_i and \mathbf{s}_j in space, and τ the time lag. The simulation method proposed here relies on the fact that the process $\{X(\mathbf{s}, t)\}$ can be obtained by transforming a parent-Gaussian process $\{Z(\mathbf{s}, t): \mathbf{s} \in \mathbb{R}^2, t \in [0, \infty)\}$ with a standard Gaussian marginal $\Phi_Z(z)$ and a suitable STCS $\rho_Z(\delta, \tau)$ that needs to be specified. One of the first papers that applied this idea was by Li and Hammond (1975) to generate univariate temporal time series using a few continuous marginals and preserving the autocorrelation up to lag 2.

This approach requires a two-fold transformation involving the marginal distribution and the STCS. Denoting the quantile function of $F_X(x)$ as Q_X , the r.v. X results from the Gaussian r.v. Z by applying the transformation $g(Z) := Q_X(\Phi_Z(Z))$. While the nonlinear transformation $g(Z(\mathbf{s}, t))$ applied to a Gaussian process yields a process with the desired marginal distribution, it affects the STCS $\rho_Z(\delta, \tau)$. Particularly, let (Z_1, Z_2) be a pair of r.v.'s following a bivariate normal distribution with correlation ρ ; then, nonlinear transformations $X_1 = g_1(Z_1)$ and $X_2 = g_2(Z_2)$ lead to a pair of r.v.'s (X_1, X_2) with correlation less (in absolute values) than ρ . The proof of this result is credited to Lancaster (1957) who referred to it as the maximal property of the bivariate normal distribution. Yet similar results were proven even earlier in the works of Maung (1941) and Gebelein (1941); the topic was also extensively studied by Rényi (1959). This result is obviously valid in a spatiotemporal framework, that is, for any bivariate vector $(Z(\mathbf{s}_i, t), Z(\mathbf{s}_j, t - \tau))$ we have $|\text{Cor}(g(Z(\mathbf{s}_i, t)), g(Z(\mathbf{s}_j, t - \tau)))| < |\text{Cor}(Z(\mathbf{s}_i, t), Z(\mathbf{s}_j, t - \tau))|$ where g is the nonlinear transformation leading to the desired marginal.

As the function g yields a target STCS $\rho_X(\delta, \tau)$ weaker than the parent-Gaussian STCS $\rho_Z(\delta, \tau)$, we need to suitably inflate the $\rho_Z(\delta, \tau)$ in order to obtain the desired $\rho_X(\delta, \tau)$. Thus, simulating the process $\{X(\mathbf{s}, t)\}$ requires the introduction of a correlation transformation function (CTF) \mathcal{T} yielding the required $\rho_Z(\delta, \tau)$ corresponding to the desired $\rho_X(\delta, \tau)$ such that $\rho_Z(\delta, \tau) = \mathcal{T}(\rho_X(\delta, \tau))$. In other words, \mathcal{T} takes as argument the target STCS and returns the inflated STCS of the parent-Gaussian process. In general, the functions linking the Gaussian and transformed correlations do not have analytical expressions apart from a few specific cases. For example, Baum (1957) provided the analytical expression when the target marginal is uniform; analytical expressions also exist for the Lognormal distribution (Matalas, 1967). In his seminal paper, Lancaster (1957) proposed the use of Hermite-Chebyshev polynomial expansions as a method to form this link; such methods have further been improved, for example, by van der Geest (1998).

Here, we build on the results shown in Papalexiou (2018) introducing a technique for fast and easy estimation of \mathcal{T} based on a simple parametric expressions of a CTF applicable also for mixed-type marginals. We provide a brief summary of the procedure to make the paper standalone. Apart from the Hermite-Chebyshev expansion proposed by Lancaster (1957) another idea uses the expectation of a transformed random variable (Feller, 1971, p. 5). If $X = g(Z)$ emerges by transforming Z and the pdf $f_Z(z)$ is known, then the expectation $E(X) = \int_{-\infty}^{\infty} g(z)f_Z(z)dz$ is calculated without the need to know the pdf of X . This is valid also in two dimensions and the correlations of the processes $\{X(\mathbf{s}, t)\}$ and $\{Z(\mathbf{s}, t)\}$ can be linked by

$$\rho_X(\delta, \tau) = \mathcal{R}(\rho_Z(\delta, \tau)) := \frac{\mathcal{C}(\rho_Z(\delta, \tau)) - \mu_X^2}{\sigma_X^2} \quad (13)$$

where μ_X and σ_X are the mean and standard deviation of X , respectively, and

$$\begin{aligned} \mathcal{C}(\rho_Z(\delta, \tau)) &:= E(X_1 X_2) = E\left(Q_X(\Phi_Z(Z_1)) Q_X(\Phi_Z(Z_2))\right) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Q_X(\Phi_Z(z_1)) Q_X(\Phi_Z(z_2)) \varphi_{Z_1 Z_2}(z_1, z_2; \rho_Z(\delta, \tau)) dz_1 dz_2 \end{aligned} \quad (14)$$

where $Z_1 = Z(\mathbf{s}_i, t)$, $Z_2 = Z(\mathbf{s}_j, t - \tau)$, $\Phi_Z(z) = (1 - \text{erf}(-z/\sqrt{2}))/2$, and $\varphi_{Z_1 Z_2}(\cdot)$ is the bivariate standard Gaussian pdf. Even though the double integral $\mathcal{C}(\rho_Z(\delta, \tau))$ does not generally have an analytical solution, its numerical estimation is quite simple and it can be performed for a fixed number of values in order to obtain a map $\rho_Z \rightarrow \rho_X$. The resulting (ρ_X, ρ_Z) pairs describing the required correlation transformation function are interpolated by fitting a simple parametric function. In particular, Papalexiou (2018) extended the use of Equation (14) by allowing the mixed-type quantiles and suggested the following function

$$\rho_Z = \mathcal{T}(\rho_X; b, c) = \frac{(1 + b\rho_X)^{1-c} - 1}{(1 + b)^{1-c} - 1} \quad (15)$$

The CTF \mathcal{T} in Eq. (15) passes from the (0,0) and (1,1) points (as for ρ_Z equal to 0 and 1, ρ_X is also 0 and 1) and is concave, fulfilling thus the condition $|\rho_Z| \geq |\rho_X|$ for $b > 0$ and $c \geq 0$. Finally, the link between the two correlations can also be established using the Jacobian of the transformation to estimate directly the bivariate density $f_{X_1 X_2}(x_1, x_2)$ and in sequence the covariance (for more details see Papalexiou, 2018).

2.3 Unified Space-time Correlation Structure (STCS)

A space-time correlation structures (STCS) $\rho(\delta, \tau; \theta)$ is a single mathematical expression that provides the correlation between two r.v.'s $X_1(t)$ and $X_2(\tau)$ located at two points of the RF and have Euclidian distance $\|\delta\|$ and are lagged in time by τ . This approach provides an elegant way to express the spatiotemporal properties of an RF using a single expression. STCS should have desirable mathematical properties (positive definiteness) that enable the generation of RFs. The literature on STCS's is vast with efforts focusing on constructing positive definite expressions. Some celebrated expressions include the Spherical, Stable, Gaussian, Whittle-Matérn, Bessel, and Cauchy (Wackernagel, 2003 pp. 334–336). This topic is still an active area of research in geostatistics with many interesting recent contributions (e.g. Alegría & Porcu, 2017; Bogaert & Christakos, 1997; Gneiting et al., 2006; Ma, 2002; Porcu et al., 2016; Schlather, 2012).

Here we propose and explore a general framework to construct potentially positive definite and flexible STCS's. For the univariate case, it was proposed that survival functions (abbreviated here as sf's and defined by $\bar{F}(x) = 1 - F(x)$ where $F(x)$ is a valid distribution function) can be used as correlation structures and several cases were demonstrated (Papalexiou, 2018). Here we extend this approach by defining an STCS as:

$$\rho(\delta, \tau; \theta) = C(\bar{F}_S(\delta; \theta_S), \bar{F}_T(\tau; \theta_T); \theta_C) \quad (16)$$

where C is a valid copula function with parameter vector θ_C ; $\bar{F}_S(\delta; \theta_S)$ and $\bar{F}_T(\tau; \theta_T)$ are survival functions with $(\delta, \tau) \in [0, \infty)$ and parameter vectors θ_S and θ_T , respectively. Note that this construction leads to functions with desired properties for an STCS, that is strictly and monotonically decreasing functions with $\rho(0, 0; \theta) = 1$ and $\lim_{\delta, \tau \rightarrow \infty} \rho(\delta, \tau; \theta) = \lim_{\delta \rightarrow \infty} \rho(\delta, \tau; \theta) = \lim_{\tau \rightarrow \infty} \rho(\delta, \tau; \theta) = 0$; it is stressed yet that solely these conditions do not suffice in general to lead in positive define STCS's. The parameter vector of the STCS is $\theta = \theta_S \cup \theta_T \cup \theta_C$. Also, the $\bar{F}_S(\delta; \theta_S)$ and $\bar{F}_T(\tau; \theta_T)$ can be interpreted as functions describing marginal correlations in space and time, respectively. Precisely, marginal space and temporal correlation structures are defined for $\tau = 0$ and $\delta = 0$, that is, $\rho(\delta, 0) = \bar{F}_S(\delta; \theta_S)$ and $\rho(0, \tau) = \bar{F}_T(\tau; \theta_T)$. We stress that the survival and copula functions are not used in the typical probabilistic sense; instead, they are used as mathematical functions enabling to build flexible strictly decreasing two-dimensional functions that could form potential spatiotemporal correlation structures.

Theoretically, this framework enables the construction of an infinite number of STCS's. It is out of the scope of this study, however, to investigate the mathematical properties and try to assess which copula families and which sf's can be combined leading to positive definite STCS's (see Section 5 for further remarks). As a general principle yet we propose the copula functions and the sf's used to be as simple as possible with analytical expressions to facilitate implementation. Note that we name the STCS's by the copula name and the corresponding distribution name of the sf used.

Based on this rationale, and as a specific example of this framework, we propose a general purpose STCS by combining the Clayton Copula

$$C(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta} \quad (17)$$

with Weibull sf's, that is, $u_1 = \bar{F}_S(\delta) = \exp(-(\delta/b_S)^{c_S})$ and $u_2 = \bar{F}_T(\tau) = \exp(-(\tau/b_T)^{c_T})$. This leads to the Clayton-Weibull STCS

$$\rho_{CW}(\delta, \tau) = \left(\exp\left(\theta \left(\frac{\delta}{b_S}\right)^{c_S}\right) + \exp\left(\theta \left(\frac{\tau}{b_T}\right)^{c_T}\right) - 1 \right)^{-\frac{1}{\theta}} \quad (18)$$

with parameter vector $(b_S, c_S, b_T, c_T, \theta)$. The parameters have the same meaning as in the individual sf's or the copula function, that is, b_S and b_T are scale parameters while c_S , c_T , and θ are shape parameters. We have tested extensively the CW model and used it in most of the applications we present here (see sections 3.1, 3.2, and 4). The Clayton-Weibull STCS was positive definite in all cases we present, yet we can also construct pathological non positive cases especially by using temporal marginal correlations that are much stronger than spatial ones. Illustrations of the Clayton-Weibull STCS in the form of contour plots are given in Figure 1a and Figure 2b.

Some other simple copula families with potential use in constructing STCS's are the Gumbel-Hougaard, Ali-Mikhail-Haq, Joe, Farlie-Gumbel-Morgenstern, among others (Nelsen, 2006; Salvadori et al., 2007). Note that potentially different sf's can be combined with a copula. For example, the marginal space correlation sf could be a Pareto II while the marginal temporal correlation could be a Weibull resulting in a Clayton-ParetoII-Weibull STCS. We stress again, however, that the complete parameter space resulting in valid positive definite STCS's needs detailed investigation; it is a complex mathematical problem that goes beyond the scopes of this study.

2.4 Space and Time Separable Correlation Structures

Equation (16) generally yields the so-called nonseparable STCS that allow for space-time interaction, which is quantified by the copula parameter(s). In particular, for specific values of the parameter(s) θ_C , most of the copula families usually applied in the hydrological literature yield (or tend to) the so-called product copula (modeling independence), which in turn corresponds to a separable STCS in the multiplicative form (on the condition that it is positive definite)

$$\rho(\delta, \tau; \theta) = \bar{F}_S(\delta; \theta_S) \cdot \bar{F}_T(\tau; \theta_T) \quad (19)$$

For the Clayton copula in Equation (17), the product copula corresponds to limiting case $\theta = 0$ (Salvadori et al., 2007; p. 237). Therefore, building STCF via copulas allow modulating the spatio-temporal dependence including independence and therefore separable STCF, if the selected copula family includes the product copula as a special case.

The hypothesis of STCS separability means that the STCS can be decomposed into the product (or sum) of a purely spatial and a purely temporal correlation functions (e.g. Gneiting, 2002). Intuitively, this hypothesis corresponds to assuming that the cross correlation (spatial correlation) can be studied independently of the autocorrelation (temporal correlation; see e.g. Genton, 2007; Gneiting et al., 2006 and references therein for a technical discussion). Of course, more general types of STCSs do exist (Genton & Kleiber, 2015; Gneiting et al., 2006), and separable STCS's have some shortcomings (Cressie & Huang, 1999, p. 1331; Kyriakidis & Journel, 1999, pp. 664–666). However, although non-separable models are often physically more realistic, separable models of space-time covariance functions are valuable in both physical and health applications (Kolovos et al., 2004). Indeed, this class of models are the simplest way to account for spatiotemporal dependence especially in cases, such as the occurrence of rare events, where there are often not enough data to justify the use of more complex dependence structures. Separable STCS's are also easier to estimate and allow for quite simple simulation.

Following the rationale introduced in Papalexiou (2018), and also used in the previous section, we suggest for $\bar{F}_S(\delta; \theta_S)$ and $\bar{F}_T(\tau; \theta_T)$ any valid survival function that can be used as a valid correlation structure. It should be clear that sf's are used as parametric functions and not as probability laws. Two simple and flexible models are the Weibull and Pareto II structures, given, respectively, by

$$\rho_W(\tau) = \exp\left(-\left(\frac{\tau}{b_T}\right)^{c_T}\right) \quad (20)$$

$$\rho_{\text{PII}}(\tau) = \left(1 + c_T \frac{\tau}{b_T}\right)^{-1/c_T} \quad (21)$$

These structures have one scale $b > 0$ and one shape $c > 0$ parameter and of course can be expressed in terms of space distance δ instead of temporal lag τ . The naming convention we follow is the same as in the previous section and uses the corresponding distribution names of the sf's used as space and temporal correlation structures. For example, a model that allows for slow decaying power-type spatial correlation and stretched exponential temporal correlations would be the separable ParetoII-Weibull STCS, given by

$$\rho_{\text{PIIW}}(\delta, \tau) = \left(1 + c_S \frac{\delta}{b_S}\right)^{-1/c_S} \cdot \exp\left(-\left(\frac{\tau}{b_T}\right)^{c_T}\right) \quad (22)$$

with parameter vector (b_S, c_S, b_T, c_T) .

Along with sf's, any other monotonically decreasing function with $\rho(0; \boldsymbol{\theta}) = 1$ and $\lim_{\tau \rightarrow \infty} \rho(\tau; \boldsymbol{\theta}) = 0$ can be used (τ can be replaced with δ), provided that such functions are positive definite (for mathematical conditions that assure positive definiteness see e.g., Cressie & Huang, 1999). For example, one could use the celebrated fGn correlation structure (Hurst, 1951; Mandelbrot & Wallis, 1968, 1969) for space and the Markovian structure for time creating the simple two parameter separable STCS

$$\rho_{\text{fGnM}}(\delta, \tau) = \frac{1}{2}(|\delta - 1|^{2H} - 2|\delta|^{2H} + |\delta + 1|^{2H}) \cdot \rho_1^\tau \quad (23)$$

where H is the Hurst coefficient aiming to quantify the strength of potential spatial persistence and ρ_1 is the lag-1 temporal autocorrelation. Specifically for one-dimensional power-type correlation structures several other formulas have been proposed (e.g. Gneiting & Schlather, 2004; Gneiting, 1999, 2000; Martin & Walker, 1997). Obviously, the separable fGn-Markovian can be extended in a non-separable framework using copulas as described in section 2.3.

In practice, using a non-separable (Equation (16)) or separable (Equation (19)) STCS to describe the empirical STCS should be investigated. If using a separable STCS results in small differences from the empirical STCS, or, the estimated copula parameter in the non-separable version indicates weak space-time interrelation, then the separable formula should be preferred. This is specifically true for fields that are strongly correlated in time. For example, if generation of Gaussian fields is made using high order Multivariate Autoregressive models (MAR), then their formulation for separable functions is heavily simplified as only diagonal coefficient matrices \mathbf{A}_i are used. Although the diagonal MAR models can explicitly preserve only the spatial structure for $\tau = 0$, and the autocorrelation structure at single field points ($\delta = 0$), this does not imply that lagged in time and space points have zero correlation; actually, the correlation between any two points and for any temporal lag τ can be estimated by the chosen separable STCS.

To further clarity, theoretical STCS's are mathematical surfaces and can be compared with their empirical counterparts. The best option is a surface that guarantees the best fit in a parsimonious way. Fitting a separable STCS that results in agreement between the spatial $\rho_X(\delta; \boldsymbol{\theta})$ and temporal $\rho_X(\tau; \boldsymbol{\theta})$ correlation structures with their empirical counterparts $\hat{\rho}_X(\delta)$ and $\hat{\rho}_X(\tau)$, does not imply agreement between the lagged in space ($\delta > 0$) and time ($\tau > 0$) correlations $\rho_X(\delta, \tau; \boldsymbol{\theta}) = \rho_X(\delta; \boldsymbol{\theta})\rho_X(\tau; \boldsymbol{\theta})$ with the empirical ones $\hat{\rho}_X(\delta, \tau; \boldsymbol{\theta})$. It should be clear, however, that using a separable STCS also does not imply zero values for the lagged space-time correlations; these values can be estimated by $\rho_X(\delta; \boldsymbol{\theta})\rho_X(\tau; \boldsymbol{\theta})$ and compared with the empirical ones. If these differences are small, then simulations using separable functions are computationally more efficient than using nonseparable functions.

2.5 Blueprint for simulating fields in simple steps

The general framework for simulating random fields having a desired marginal probability distribution (including mixed-type and discrete marginals) and spatiotemporal correlation structure is summarized in the steps:

1. Fit or choose an appropriate target marginal distribution $F_X(x)$ that describes the values of the RF's (section 2.1).
2. Estimate the correlation transformation function $\mathcal{T}(\rho_X; b, c)$ for the target marginal distribution (section 2.2).
3. Describe the target spatiotemporal correlation structure (STCS) of the field by using nonseparable (section 2.3) or separable (section 2.4) STCS $\rho_X(\delta, \tau)$. Separable STCS's facilitate computation but should be used with caution.
4. Transform the target STCS by using the correlation transformation function to estimate the Gaussian STCS, that is, $\rho_Z(\delta, \tau) = \mathcal{T}(\rho_X(\delta, \tau); b, c)$.
5. Generate Gaussian RFs by using multivariate AR(p) models (or any other method) that reproduce the parent-Gaussian STCS.
6. Transform the Gaussian RFs using the marginal-back transformation.

We emphasize that we have not observed any effect in the positive definiteness of an STCS when it was transformed by the correlation transformation function \mathcal{T} . Yet Steps 3 and 4 can be modified to assure the Gaussian STCS is positive definite. Particularly, once the correlation transformation function $\mathcal{T}(\rho_X; b, c)$ is defined (Step 2), instead of fitting a positive definite function $\rho_X(\delta, \tau)$ to the observed field, we can fit a function of the form $\rho_X(\delta, \tau) = \mathcal{T}^{-1}(\rho_Z(\delta, \tau))$; where \mathcal{T}^{-1} is the inverse of the correlation transformation function and $\rho_Z(\delta, \tau)$ can be any mathematically proven positive definite STCS. This assures the Gaussian field will be generated by using the positive definite STCS $\rho_Z(\delta, \tau)$.

We note that this framework can be readily extended to reproduce seasonality by switching the model that generates the Gaussian fields based on a seasonal cycle and using the corresponding seasonal marginal distributions (for details on the univariate case the reader is referred to Papalexiou, (2018)).

3. Random Fields in Action

3.1 Precipitation

We generate precipitation RFs with a mixed-type marginal distribution having probability zero $p_0 = 0.8$ and with nonzero values following a J-shaped power-type distribution, that is, the Burr type XII $\mathcal{BrXII}(3,0.9,0.2)$. This choice is consistent with global analyses on daily precipitation showing that the Generalized Gamma and the Burr type XII are two effective models to describe precipitation (Papalexiou & Koutsoyiannis, 2012, 2016). We describe the spatiotemporal correlation structure by the Clayton-Weibull model $\rho_{CW}(\delta, \tau)$, given in Equation (18), with parameters $(b_S, c_S, b_T, c_T, \theta) = (20, 0.7, 1.1, 0.8, 2)$ (Figure 1a). For the target p_0 and marginal \mathcal{BrXII} distribution the correlation transformation function is $\mathcal{T}(\rho_X; 13.2, 0.75)$ (Figure 1b) which is used to transform the target STCS into the Gaussian STCS (Figure 1c).

Note that the target STCS with these parameters (Figure 1a) provides a stretched-exponential decaying spatial correlation structure (for example for temporal lag $\tau = 0$ and space distance $\delta = 1$ and $\delta = 100$ we have $\rho_{CW}(1, 0) = 0.88$ and $\rho_{CW}(100, 0) = 0.05$, respectively). The temporal structure is assumed to be less intense, as is typically observed in daily precipitation time series, yet decaying slower than a Markovian structure; for example the autocorrelation of each point for lag $\tau = 1$ and $\tau = 5$ is $\rho_{CW}(0, 1) = 0.4$ and $\rho_{CW}(0, 5) = 0.03$, respectively. For demonstration we use the full MAR(5) that reproduces precisely the spatiotemporal structure up to lag $\tau = 5$ and up to any distance

δ . Finally, we generate Gaussian RFs from the MAR(5) and transform them to the target fields using the mixed-type quantile function (see also Papalexiou, 2018). A sample sequence of nine RFs (Figure 1d) demonstrates characteristic properties of precipitation fields, that is, the strong spatial structure that yields totally dry or wet fields but also partly dry. The temporal structure is better visualized in a sequence of 80 RFs (Figure A1) where a strong clustering is apparent with several wet days followed by several dry days.

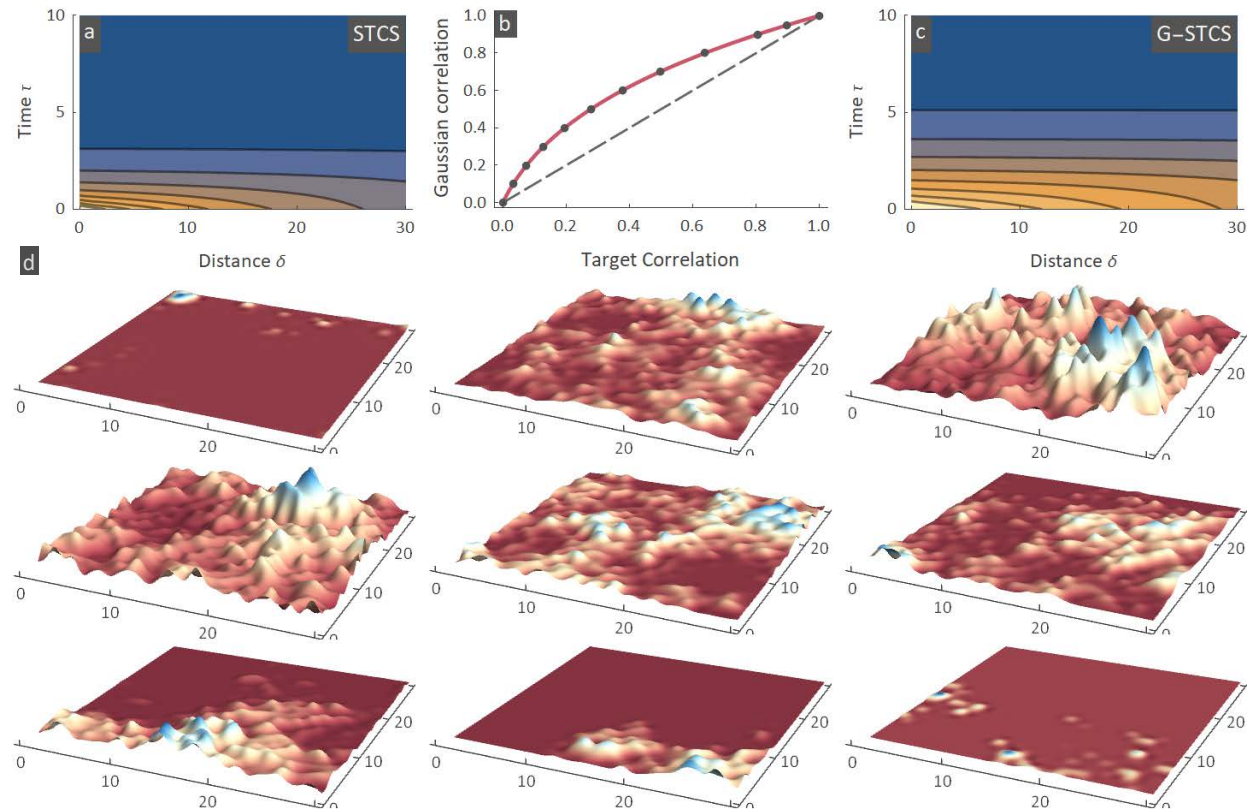


Figure 1. Precipitation fields simulation. (a) The target spatiotemporal correlation structure (STCS). (b) The correlation transformation function for target marginal. (c) The transform Gaussian STCS. (d) Sequence of nine simulated precipitation fields with target marginal distribution the Burr type XII $\mathcal{BrXII}(3,0.9,0.2)$ and probability zero $p_0 = 0.8$ (see Figure A1 for a sequence of 80 RFs).

We note that stochastic modelling of precipitation has been evolving over several decades and many methods exist dating back to the 70's (see e.g., Amorcho & Wu, 1977; Mejía & Rodríguez-Iturbe, 1974) and 80's (see e.g., Bell, 1987). Apart from methods based on Gaussian parent processes, many more exist based on different rationales. For example, there are methods using atmospheric circulation patterns (e.g., Bardossy & Plate, 1992), wavelets transforms (Kumar & Foufoula-Georgiou, 1997), or, are based on fractal and multifractal techniques (see e.g., Foufoula-Georgiou, 1998; Gupta & Waymire, 1993; Langousis et al., 2009; S. Lovejoy & Mandelbrot, 1985; Shaun Lovejoy & Schertzer, 1990; Schertzer & Lovejoy, 1987; Veneziano et al., 2006). Additionally, models for high resolution space-time rainfall modeling (e.g., at 5 min, and 1 km²) may need to include advection and anisotropy components in order to simulate fields more realistically at these scales. For example, Jinno et al. (1993) and Kawamura et al. (1996) use advection-diffusion equations, Pegram and Clothier (2001) introduced the "String of Beads" method, Mantoglou and Wilson (1982) investigated the "Turning bands method" (later modified by Lebois and Creutin (2013)); many more simulators

can also be found in the literature (see e.g., Baxevani & Lennartsson, 2015; Benoit et al., 2018; Paschalis et al., 2013; Peleg et al., 2017). Our technique can be easily modified to include anisotropy (see Figure A2) and advection, but such details are out of the scope of this study. Yet the typical approach for introducing anisotropy is trivial and is based on simple coordinate transformations (see e.g., Youngman & Stephenson, 2016). We provide these equations in the Annex A and we show a simulation with anisotropy for a sequence of 80 RFs (see Figure A2).

3.2 Relative humidity

In this case we generate relative humidity RFs. As a marginal distribution we use a characteristic model, that is, the Beta distribution $\mathcal{B}(\gamma_1, \gamma_2)$ with probability density function (pdf)

$$f_{\mathcal{B}}(x) = \frac{x^{\gamma_1-1}(1-x)^{\gamma_2-1}}{\mathcal{B}(\gamma_1, \gamma_2)} \quad (24)$$

with $\mathcal{B}(\gamma_1, \gamma_2) = \int_0^1 x^{\gamma_1-1}(1-x)^{\gamma_2-1}dx$. The $\mathcal{B}(\gamma_1, \gamma_2)$ distribution is defined in $[0,1]$, and thus, can describe variables bounded in this range such as relative humidity. Of course, any other consistent model can be used; another option is the Kumaraswamy distribution $\mathcal{Ku}(\gamma_1, \gamma_2)$ which has simpler and analytical cdf and quantile functions compared to the Beta distribution. Here we use the $\mathcal{B}(7,1.5)$ as a target marginal (Figure 2a) and a very strong spatiotemporal structure, that is, the Clayton-Weibull STCS $\rho_{\text{CW}}(\delta, \tau)$ with parameters $(\beta_S, \gamma_S, \beta_T, \gamma_T, \theta) = (100, 1, 2.8, 1, 0.5)$ (Figure 2b). The spatial structure for $\tau = 0$ is a very strong Markovian, that is, $\rho_{\text{CW}}(\delta, 0) \approx 0.99^\delta$, while the temporal structure for $\delta = 0$ is also Markovian, less intense though, with $\rho_{\text{CW}}(0, \tau) \approx 0.70^\tau$.

For the target $\mathcal{B}(7,1.5)$ distribution the estimated correlation transformation function $\mathcal{T}(\rho_X; 0.51, 0.35)$, in contrast with the precipitation case in the previous demonstration, shows that only a slight correlation inflation is needed. This is apparent by comparing the target STCS (Figure 2b) with the transform Gaussian STCS (Figure 2c) that are very similar. For more details on which distributions result in highly concave correlation transformation functions see the Section 3.3 in Papalexiou (2018). Here we use the MAR(1) to generate Gaussian RFs and transform them to the target fields using the quantile function of the $\mathcal{B}(7,1.5)$ distribution. A random sequence of nine RFs (Figure 2d) shows the very strong spatial structure with smooth spatial changes or even with fields preserving very similar values for the whole field area. The temporal clustering is better demonstrated by a longer random sequence of RFs (Figure A3) where we observe consecutive fields with high RH values alternating with fields having low RH values.

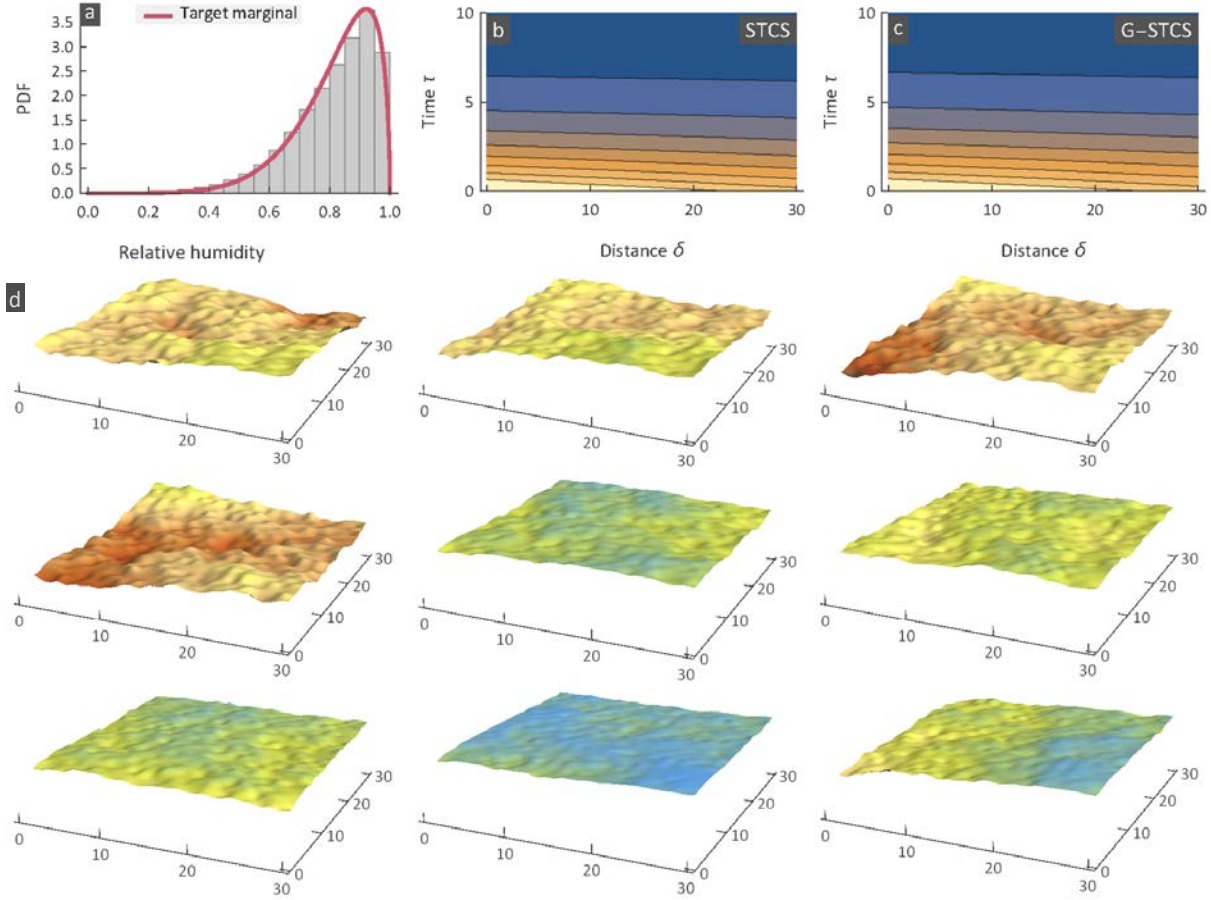


Figure 2. Simulation of relative humidity (RH) fields. (a) The target probability density function of the RH fields. (b) The target spatiotemporal correlation structure (STCS), and (c) the transformed Gaussian STCS. (d) Sequence of nine simulated relative humidity fields (see Figure A3 for a sequence of 80 RFs).

3.3 Daily maximum temperature fields

Here we demonstrate the generation of spatiotemporal fields using independent space and time correlation structures; particularly we use the separable ParetoII-Weibull STCS in Equation (22). We generate fields that could resemble daily maximum temperature. Since we are assuming fields of daily maxima we use as target marginal the Gumbel distribution $\mathcal{G}u(\alpha, \beta)$ with distribution function (cdf)

$$F_{\mathcal{G}u}(x) = \exp\left(-\exp\left(-\frac{x-\alpha}{\beta}\right)\right) \quad (25)$$

It is well-known that such a model can represent maxima emerging from exponential-type parent distributions as it is expected for temperature. Particularly, we use the $\mathcal{G}u(23,1)$ assuming that the field is isotropic with mean value $\mu \approx 23.6^\circ\text{C}$ and standard deviation $\sigma \approx 1.3^\circ\text{C}$ (Figure 3a).

The flexibility of this approach allows the choice of independent correlation structures for space and time, having the drawback yet of not being able to preserve explicitly the temporal correlation of distant points in space. As target spatial correlation structure we use a slowly decaying Pareto II correlation structure (Equation (20)), that is, $\rho_{\text{SCS}}(\delta) = \rho_{\text{PII}}(\delta; 10,5)$ (red line in Figure 3c). For temporal correlation structure, representing the autocorrelation of each pixel, we use the Weibull structure in Equation (21) with $\rho_{\text{TCS}}(\tau) = \rho_{\text{W}}(\tau; 4,0.75)$ (blue line; Figure 3c). The fitted correlation

transformation function $\mathcal{T}(\rho_X; 0.005, 20.5)$ (Figure 3b) for the target $\mathcal{G}u$ distribution shows that inflation is negligible.

To reproduce the spatial and temporal correlation of each pixel we use a high order MAR(20) model with diagonal matrices. As previously, we transform the Gaussian fields using the quantile function of the target $\mathcal{G}u$ distribution. This approach facilitates the computations and the model reproduces precisely the spatial correlations for time lag $\tau = 0$ (red dots in Figure 3c), and the temporal autocorrelation of each pixel up to lag $\tau = 20$ (blue dots in Figure 3c). A random sequence of nine daily T_{\max} RFs (Figure 2d) illustrates the strong spatial structure but also of the temporal clustering of high and low value fields. The latter is clearly illustrated in a longer random sequence of 80 RFs (Figure A3).

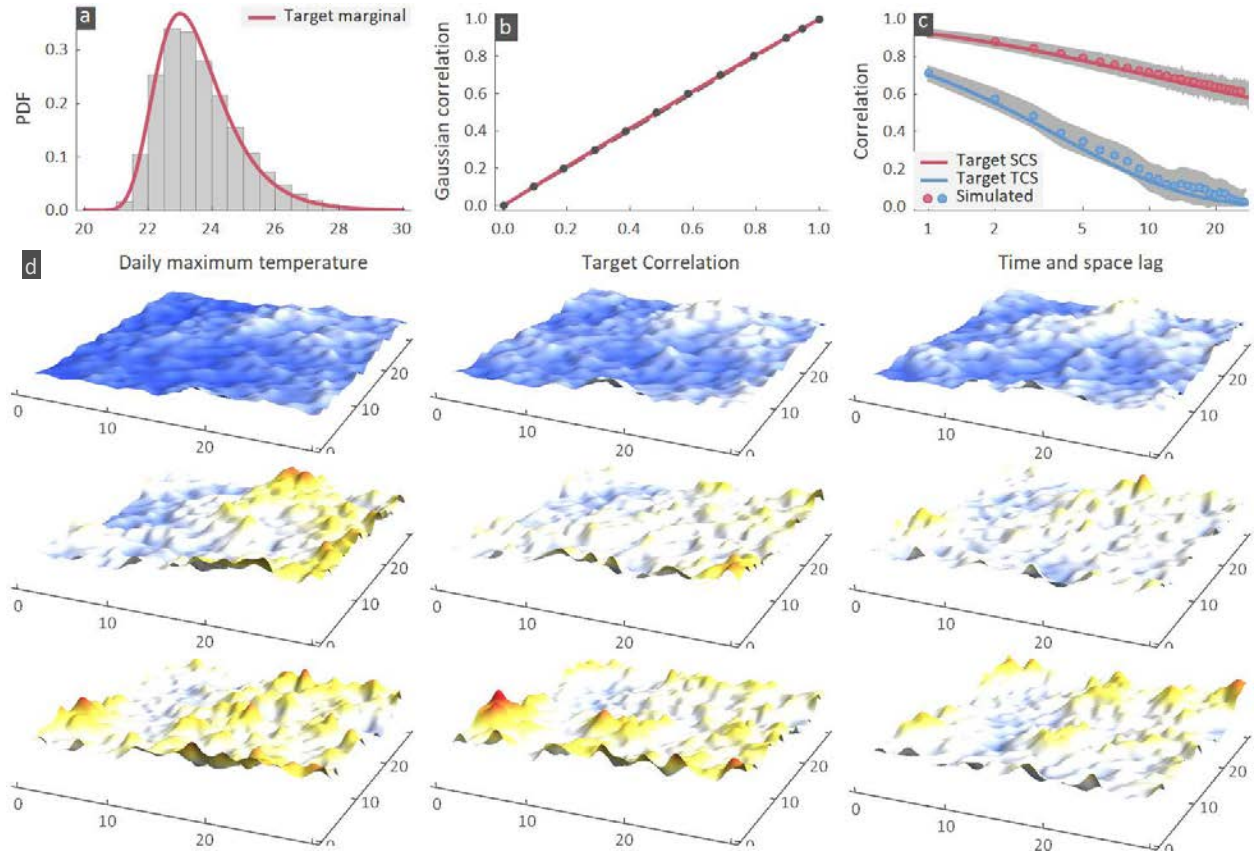


Figure 3. Daily maximum temperature (T_{\max}) fields simulation using separable space-time correlation structures. (a) The target probability density function of the T_{\max} fields. (b) The correlation transformation function for target marginal. (c) The target space (SCS) and temporal (TCS) correlation structures compared with the simulated ones. (d) Sequence of nine simulated T_{\max} with target marginal distribution the Gumbel $\mathcal{G}u(23,1)$ (see Figure A4 for a sequence of 80 RFs)

4. Real-world Simulation of Daily Precipitation Fields

Here we explore the properties of real-world precipitation fields and we simulate synthetic ones. We extracted daily precipitation fields from the CHIRPS (Climate Hazards Group InfraRed Precipitation with Station data) dataset for a U.S. region defined by latitude 39.475° to 40.925° and longitude -111.425° to -109.925° . The spatial resolution of the data is 0.05° and cover the 1981-2017 period (see Figure 5 for six observed fields and Figure A5 for a sequence of 80 observed fields).

The empirical fields are represented by 30×30 grid cells and we assume that the field is described by the same marginal distribution, that is, the distribution of the nonzero values as well as the probability of zero values of each cell is the same. The empirical properties showed that the distribution on nonzero value is bell-shaped with a light tail. Flexible and consistent distributions for point measurements, such as the Generalized Gamma and the Burr type XII (Papalexiou & Koutsoyiannis, 2012), did not provide an adequate fit. This led us to introduce a new flexible 3-parameter distribution based on rationale described in Section 2.1. We name this distribution the Generalized Standard Gompertz $\mathcal{GSG}(\beta, \gamma_1, \gamma_2)$; it has an analytical cdf given by

$$F_{\mathcal{GSG}}(x) = 1 - \exp\left(1 - \exp\left(\left(1 + \left(\frac{x}{\beta}\right)^{\gamma_1}\right)^{\gamma_2} - 1\right)\right) \quad (26)$$

For $\gamma_1 = \gamma_2 = 1$ the cdf simplifies to $F_{\mathcal{GSG}}(x) = 1 - \exp(1 - \exp(x/\beta))$ which can be considered as the “standard” Gompertz distribution emerging for shape parameter equal to 1 (Garg et al., 1970; Gompertz, 1815). The $\mathcal{GSG}(\beta, \gamma_1, \gamma_2)$ is defined in $[0, \infty)$ and has one scale parameter $\beta > 0$ and two shape parameters $\gamma_1 > 0$ and $\gamma_2 > 0$. Its pdf can be bell- or J-shaped for $\gamma_2 < 1$ and $\gamma_2 > 1$, respectively, while the heaviness of the tail is controlled by γ_2 (larger γ_2 lead to heavier tails).

For demonstration we use the precipitation fields of January. The empirical analysis shows that the $\mathcal{GSG}(0.99, 2.73, 0.12)$ provides a good fit to the observed data (Figure 4a) while the probability dry is $p_0 = 0.78$. The distribution was fitted by minimizing the square error between the observed values and the predicted ones by the quantile of the \mathcal{GSG} . To describe the spatiotemporal correlation structure of the fields, we use the Clayton-Weibull STCS introduced in Equation (18) as follows: (1) we estimated the correlations between the time series in 62,500 random grid cell pairs (the total number of pairs is too large, i.e., 30^4) and for time lags up to 2 days; (2) we stratified the estimated correlations based on the distance of the pairs and calculated the mean correlation value; (3) we fitted the Clayton-Weibull STCS by numerically minimizing the square error between observed values and predict ones by the STCS. The estimated parameter vector of the Clayton-Weibull STCS is $(\beta_S, \gamma_S, \beta_T, \gamma_T, \theta) = (3.12, 0.74, 0.44, 0.49, 0.64)$ and as Figure 4b attests the structure provides an accurate description of the observed spatiotemporal correlation values.

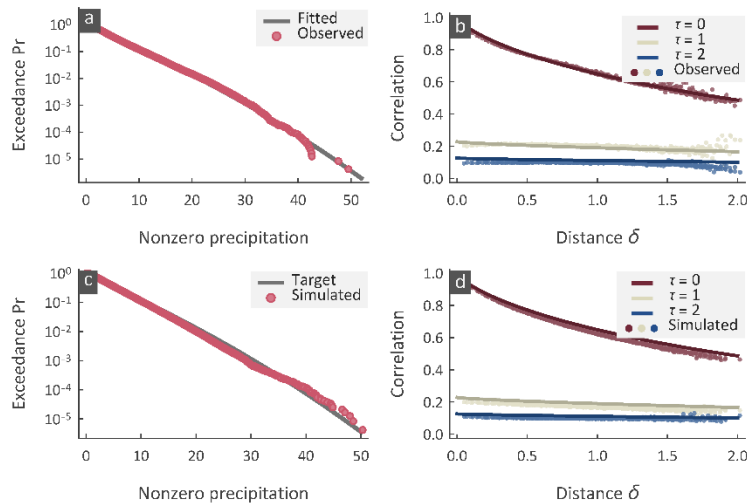


Figure 4. Assessing a real-world simulation of daily precipitation fields: (a) The fitted \mathcal{GSG} distribution to the observed nonzero precipitation. (b) The fitted Clayton-Weibull STCS compared with observed values for several time lags. (c) The empirical distribution of the generated field values compared with the target \mathcal{GSG} . (d) The empirical STCS of the simulated field compared with the target Clayton-Weibull STCS.

To simulate synthetic fields we follow the rest of the steps described in section 2.3: (1) we estimate the correlation transformation function for the fitted (target) \mathcal{GSG} and the estimated p_0 , that is, $\mathcal{T}(\rho_X; 4.8, 0.77)$; (2) transform the fitted Clayton-Weibull STCS $\rho_{CW}(\delta, \tau)$ to obtain the Gaussian STCS, that is $\rho_Z(\delta, \tau) = \mathcal{T}(\rho_{CW}(\delta, \tau); 4.8, 0.77)$; (3) use the estimated Gaussian STCS $\rho_Z(\delta, \tau)$ to estimate the parameters of a MAR(2), and (4) transform the generated Gaussian fields using the quantile of the target mixed-type marginal distribution.

We show six simulated fields in Figure 5 while a sequence of 80 simulated ones is given in Figure A6. First, as anticipated the simulated marginal distribution (Figure 4c) and the spatiotemporal structure (Figure 4d) match the target ones. Second, the comparison of the observed and simulated field (Figure 5, Figure A5, Figure A6) shows a clear spatial and temporal clustering. Sequences of no precipitation are followed by wet days with precipitating in the whole region or in parts of it. This behavior is recognized clearly also in the observed fields.

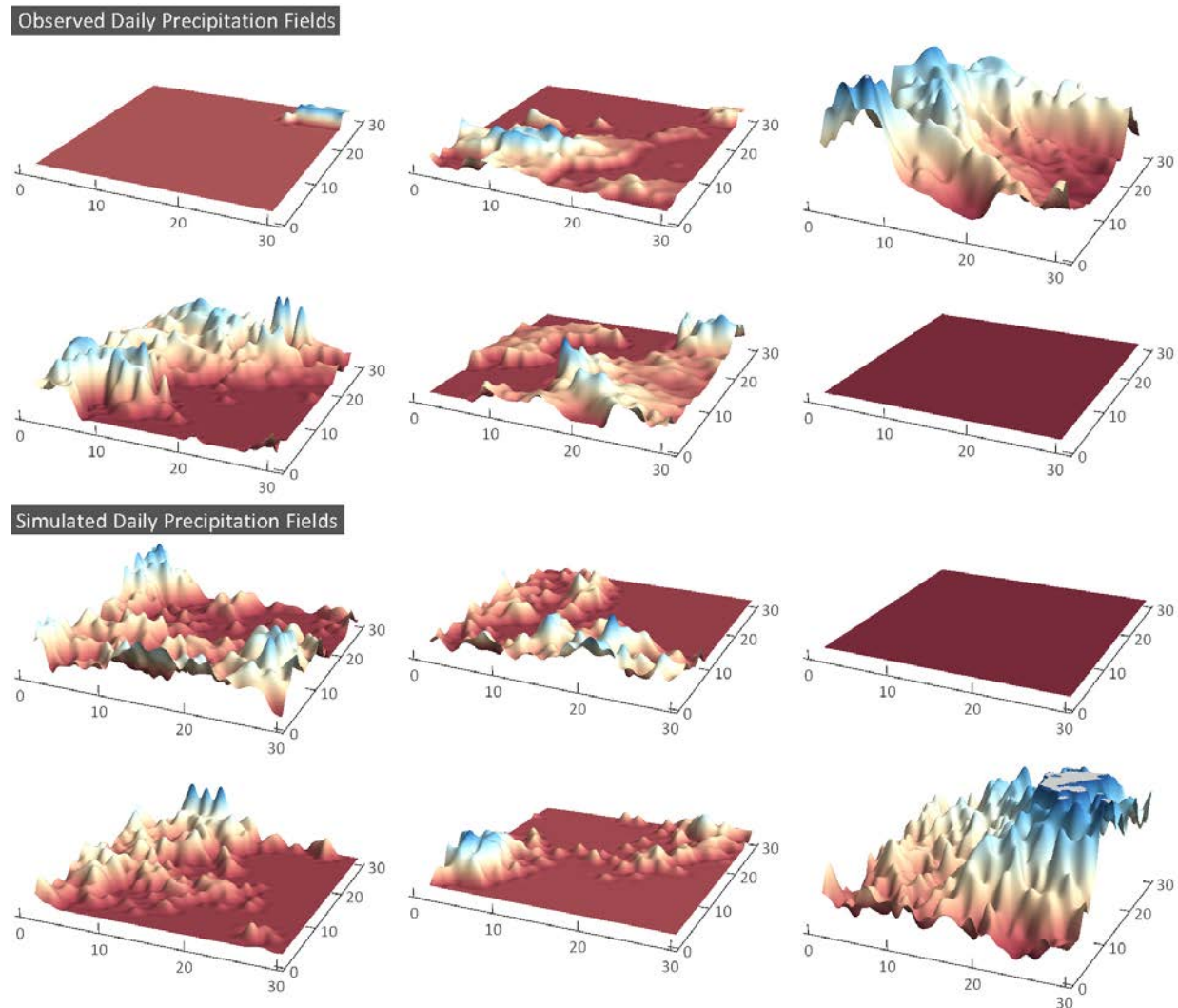


Figure 5. A sequence of six observed and six simulated precipitation fields.

5. Discussion

Several issues of this blueprint should be emphasized while others deserve further discussion. For the marginal distributions of positive variables, we considered a set of distribution families (i.e.,

Generalized Gamma, Burr type III and XII and Beta of the Second Kind) whose justification is based on simple criteria fulfilling basic distributional properties of the process under study (see Section 2.1). These families comprise several well-known and widely used distributions (Gamma, Weibull, Pareto, etc.) as special cases, but are more general and flexible (enabling the modelling of both exponential and power-type tails) while retaining parsimony in terms of number of parameters. Extending these continuous distributions by inflating them at zero we built mixed-type distributions that can describe zero-inflated variables such as rainfall, or over-threshold processes. Alternatively, these distributions can be easily combined with a Bernoulli process. We note that our framework also allows for the use of double bounded distributions such as beta with or without probability masses at one or both boundaries, describing for instance the occurrence of boundary conditions such as empty and full reservoirs.

In the proposed framework, these univariate distributions are used to transform the marginal distribution of spatiotemporal Gaussian random fields into the desired form (e.g., Generalized Gamma, Burr type XII, etc.). The spatiotemporal correlation function of these fields is suitably inflated to obtain the desired correlation after marginal transformation. To accomplish this task, we propose an efficient technique by using a transformation correlation function that enables the easy and straightforward estimation of the parent Gaussian correlation. Such an approach avoids multiple computation (or iterative procedures) of the double integral linking the correlation values of the parent and target processes, as well the use of approximating polynomials, which are often proposed in the literature to deal with the transformation procedure.

Concerning the choice of the spatiotemporal correlation function, we proposed a new method to construct models resulting from the combination of copulas and survival functions. Of course, we are aware that new spatiotemporal correlation functions should be introduced with great care in order to guarantee that they are permissible, i.e. valid functions fulfilling the condition of positive definiteness (Armstrong & Jabin, 1981). Although the proposed models have been tested over a variety of parameter sets, we did not provide theoretical proof of their positive definiteness. Moreover, some combinations of parameters surely do not fulfil this property. However, we recall that the simulation algorithm is based on vector AR models that are fed by correlation matrices resulting from the spatiotemporal correlation function. Therefore, from a practical point of view, positive definiteness of the feeding correlation matrices can be checked numerically for each specific application by calculating the sign and magnitude of their eigenvalues. Furthermore, if these matrices are not positive definite they can be corrected to fulfil this property (Higham, 2002). Yet in all our examples, the proposed spatiotemporal correlation structures were always positive definite. This can help circumventing the problem in practical applications, while further research is required to provide mathematical proofs. We also note that the spatiotemporal correlation functions can involve the space and time separable models when the copula parameter tends to the value corresponding to the so-called product copula, describing independence.

Noteworthy, the Gaussian dependence structure is characterized by zero tail dependence coefficient. Nonetheless, for finite size samples, the extreme values resulting by Gaussian copulas tend to preserve correlation values smaller than but not too far from those of models characterized by positive upper tail dependence, such as Student or extreme value (EV) copulas (e.g., Serinaldi et al., 2015). This behavior is due to the (elliptical) topology of the dependence structure that yields upper tail clustering weaker than that of EV or Student copulas but stronger, for example, than that of Frank or Clayton copulas. In this respect, a future line of research will be extending our model by including Student or grouped Student dependence structures building on Serinaldi and Kilsby (2016), thus accounting for asymptotic tail dependence. We also recall that the Gaussian dependence

structures allow for modeling second order joint properties, while higher-order joint properties can be of interest, yet they are hard to recognize and estimate based on data.

A point of special interest is the reproduction of extremes in terms of marginal properties. The frequency and the magnitude of extremes is quantified by the heaviness of the tail. Specifically, for precipitation extremes, recent global and large scale studies show that heavy-tailed distributions represent extremes more accurately (see e.g., Nerantzaki & Papalexiou, 2019; Papalexiou, AghaKouchak, et al., 2018; Serinaldi & Kilsby, 2014). In this respect, the flexibility offered by this scheme, allowing the choice of marginals with any asymptotic tail behavior (power-type, stretched exponential, lognormal, etc.), facilitates the accurate reproduction of extremes if the selected marginal is indeed an adequate model. Of course, if the focus is specifically on reproducing extremes, extreme value theory also provides other options such as max-stable processes and corresponding random fields (e.g. Blanchet & Davison, 2011; Cooley et al., 2006; Davison et al., 2012; Davison & Gholamrezaee, 2012; Robert, 2013; Martin Schlather, 2002; Tyralis & Langousis, 2019). Also, as a cautionary note, we mention that reproduction of extremes does not only rely on competent models, but also on precise data. This issue is specifically true when dealing with annual maxima where discretization bias may lead to severely underestimate the actual values. This important issue is typically neglected while it could be tackled using the “forgotten” Hershfield factor (van Montfort, 1990; Papalexiou et al., 2016).

Also, the scheme presented here guarantees the reproduction of any marginal distribution but is not designed to reproduce specific product moments as in multifractals methods (see e.g., Foufoula-Georgiou, 1998; Langousis & Veneziano, 2007; Lovejoy & Mandelbrot, 1985; Schertzer & Lovejoy, 1987). For example, if a three-parameter target marginal is fitted on the first three moments (any other fitting methods can be used) then the first three moments are explicitly preserved. All higher finite moments are completely specified by the fitted distribution. However, while the fitted marginal does not guarantee the reproduction of higher order moments, vice versa, preserving higher order moments does not guarantee the reproduction of the marginal. Similarly, this modelling approach does not explicitly reproduce moments at multiple scales; the interested reader is referred to the multifractal literature for such methods. Yet if the process at the fine scale (e.g., daily) is described with precision by the fitted model (correlation structure and marginal), then it is anticipated the aggregated process at higher scales to be consistent. Additionally, if long-term persistence is needed, it can be achieved with two ways, that is, either by selecting a very strong temporal correlation structure (see examples in Papalexiou, 2018), or by disaggregating, for example, annual totals with persistence using the spatiotemporal extension of the DiPMaC disaggregation scheme (Papalexiou, Markonis, et al., 2018). In any case, models reproducing long-term persistence should be used with caution as it might be difficult to distinguish or quantify persistence based on data (Markonis et al., 2018).

Finally, the use of multivariate AR models to simulate the parent Gaussian random fields is known to suffer from the so-called curse of dimensionality, preventing the simulation of fields over spatiotemporal lattices with many grid nodes (at least in the nonseparable case and when the MAR order is very large). Therefore, further research is required to implement more efficient generators based for instance on spectral techniques (Lang & Potthoff, 2011).

6. Summarizing and Concluding

Geophysical and environmental processes naturally evolve in space and time and their complexity usually prevents pure deterministic representations. Therefore, stochastic modelling coupled or not with deterministic components is a viable option to describe such natural processes. Despite the long history of geostatistics, modelling and simulation of random fields, that is, stochastic processes

defined over the Cartesian product of multidimensional spatial domain and unidimensional time domain, is still an active research area. First, this is due to the challenging problems posed by the above-mentioned complexity; and second, due to emerging advances in remote sensing and information technology that allow high-resolution spatiotemporal observations and efficient storage and processing. The latter is verified by the many gridded products of precipitation, temperature, etc., that have been developed over the last decade and are readily available.

In this context, we aim to introduce a blueprint for effective, efficient, and easy-to-apply simulation of spatiotemporal random fields. The generated fields preserve important characteristics such as any valid spatiotemporal correlation structure, and any prescribed marginal distribution. This includes also discrete-continuous marginals to allow for simulation of intermittency, or, of more general cases where discrete probabilities exist in more than one value, but also purely discrete marginals. The scheme uses a basic idea, dating back in the 70's, which forms a desired process by transforming a Gaussian one (parent), and thus, it is based on the Gaussian dependence topologies. We build on the methodological techniques presented in Papalexiou (2018) but our results can be extended to non-Gaussian dependences following the results shown by Serinaldi and Kilsby (2016).

As a proof of concept of the versatility of the proposed generator, we showed simulated fields for several conceptual cases mimicking zero-inflated precipitation, double-bounded relative humidity, and strongly dependent temperature maxima, as well as a real-word example concerning intermittent precipitation fields. In all cases, we showed the faithful reproduction of the desired properties. Although the current study focuses on hydroclimatic random fields evolving typically in space and time, this scheme has potential applications far beyond hydroclimatology.

Further improvements concern the introduction of spatial anisotropy beyond simple coordinate transformations, and the efficient transport/advection terms allowing the simulation of weather systems, or spatial patterns, travelling over a given region. Finally, the current scheme forms the basis to extend the DiPMaC (Disaggregation Preserving Marginal and Correlations) algorithm (Papalexiou, Markonis, et al., 2018) into spatial and spatiotemporal cases allowing downscaling that preserves the marginal distribution and the spatiotemporal correlation structure at any desired high-resolution spatiotemporal scale.

Acknowledgements

SMP was funded by the Global Water Futures program. FS acknowledges the support from the Willis Research Network. We thank three anonymous reviewers for their constructive remarks and criticisms that helped improve the original manuscript.

The observed data used are freely available at <https://www.chc.ucsb.edu/data/chirps#>. We have used partly the CoSMoS R-package available at CRAN.

Competing interests

The authors declare no competing interests.

References

- Adler, R. J., & Taylor, J. E. (2007). *Random Fields and Geometry*. New York: Springer-Verlag. Retrieved from <https://www.springer.com/it/book/9780387481128>
- Alegría, A., & Porcu, E. (2017). The dimple problem related to space–time modeling under the Lagrangian framework. *Journal of Multivariate Analysis*, 162, 110–121. <https://doi.org/10.1016/j.jmva.2017.09.001>
- Allard, D. (2012). Modeling Spatial and Spatio-Temporal Non Gaussian Processes. In E. Porcu, J. Montero, & M. Schlather (Eds.), *Advances and Challenges in Space-time Modelling of Natural Events* (pp. 141–164). Springer Berlin Heidelberg.
- Amorocho, J., & Wu, B. (1977). Mathematical models for the simulation of cyclonic storm sequences and precipitation fields. *Journal of Hydrology*, 32(3), 329–345. [https://doi.org/10.1016/0022-1694\(77\)90025-7](https://doi.org/10.1016/0022-1694(77)90025-7)

- Angulo, J., Yu, H.-L., Langousis, A., Kolovos, A., Wang, J., Madrid, A. E., & Christakos, G. (2013). Spatiotemporal Infectious Disease Modeling: A BME-SIR Approach. *PLoS ONE*, *8*(9). <https://doi.org/10.1371/journal.pone.0072168>
- Armstrong, M., & Jabin, R. (1981). Variogram models must be positive-definite. *Journal of the International Association for Mathematical Geology*, *13*(5), 455–459. <https://doi.org/10.1007/BF01079648>
- Bacro, J.-N., & Gaetan, C. (2012). A Review on Spatial Extreme Modelling. In E. Porcu, J. Montero, & M. Schlather (Eds.), *Advances and Challenges in Space-time Modelling of Natural Events* (pp. 103–124). Springer Berlin Heidelberg.
- Bardossy, A., & Plate, E. J. (1992). Space-time model for daily rainfall using atmospheric circulation patterns. *Water Resources Research*, *28*(5), 1247–1259. <https://doi.org/10.1029/91WR02589>
- Baum, R. (1957). The correlation function of smoothly limited Gaussian noise. *IRE Transactions on Information Theory*, *3*(3), 193–197. <https://doi.org/10.1109/TIT.1957.1057415>
- Baxevani, A., & Lennartsson, J. (2015). A spatiotemporal precipitation generator based on a censored latent Gaussian field. *Water Resources Research*, *51*(6), 4338–4358. <https://doi.org/10.1002/2014WR016455>
- Bell, T. L. (1987). A space-time stochastic model of rainfall for satellite remote-sensing studies. *Journal of Geophysical Research: Atmospheres*, *92*(D8), 9631–9643. <https://doi.org/10.1029/JD092iD08p09631>
- Benoit, L., Allard, D., & Mariethoz, G. (2018). Stochastic Rainfall Modeling at Sub-kilometer Scale. *Water Resources Research*, *54*(6), 4108–4130. <https://doi.org/10.1029/2018WR022817>
- Blanchet, J., & Davison, A. C. (2011). Spatial modeling of extreme snow depth. *The Annals of Applied Statistics*, *5*(3), 1699–1725. <https://doi.org/10.1214/11-AOAS464>
- Bogaert, P., & Christakos, G. (1997). Spatiotemporal analysis and processing of thermometric data over Belgium. *Journal of Geophysical Research: Atmospheres*, *102*(D22), 25831–25846. <https://doi.org/10.1029/97JD01809>
- Burr, I. W. (1942). Cumulative Frequency Functions. *The Annals of Mathematical Statistics*, *13*(2), 215–232.
- Chilès, J.-P., & Delfiner, P. (2009). *Geostatistics: Modeling Spatial Uncertainty*. John Wiley & Sons.
- Christakos, G. (2005). *Random Field Models in Earth Sciences* (Dover Ed edition). Mineola, N.Y: Dover Publications.
- Christakos, G., & Hristopulos, D. T. (1998). Environmental Exposure Fields and Their Health Effects. In G. Christakos & D. T. Hristopulos (Eds.), *Spatiotemporal Environmental Health Modelling: A Tractatus Stochasticus* (pp. 37–79). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4757-2811-8_2
- Cooley, D., Naveau, P., & Poncet, P. (2006). Variograms for spatial max-stable random fields. In P. Bertail, P. Soulier, & P. Doukhan (Eds.), *Dependence in Probability and Statistics* (pp. 373–390). New York, NY: Springer. https://doi.org/10.1007/0-387-36062-X_17
- Cressie, N., & Huang, H.-C. (1999). Classes of Nonseparable, Spatio-Temporal Stationary Covariance Functions. *Journal of the American Statistical Association*, *94*(448), 1330–1340. <https://doi.org/10.2307/2669946>
- Davison, A. C., & Gholamrezaee, M. M. (2012). Geostatistics of extremes. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *468*(2138), 581–608. <https://doi.org/10.1098/rspa.2011.0412>
- Davison, A. C., Padoan, S. A., & Ribatet, M. (2012). Statistical Modeling of Spatial Extremes. *Statistical Science*, *27*(2), 161–186. <https://doi.org/10.1214/11-STS376>
- Evstigneev, I. V., & Zhitlukhin, M. V. (2013). Controlled random fields, von Neumann–Gale dynamics and multimarket hedging with risk. *Stochastics*, *85*(4), 652–666. <https://doi.org/10.1080/17442508.2013.795565>
- Fei, X., Xiao, R., Christakos, G., Langousis, A., Ren, Z., Tian, Y., & Lv, X. (2019). Comprehensive assessment and source apportionment of heavy metals in Shanghai agricultural soils with different fertility levels. *Ecological Indicators*, *106*, 105508. <https://doi.org/10.1016/j.ecolind.2019.105508>
- Feller, W. (1971). Introduction to the Theory of Probability and its Applications, Vol. 2. *II (2. Ed.)* New York: Wiley.
- Foufoula-Georgiou, E. (1998). On scaling theories of space-time rainfall: Some recent results and open problems. *Stochastic Methods in Hydrology: Rainfall, Landforms and Floods*, 25–72.
- Garg, M. L., Rao, B. R., & Redmond, C. K. (1970). Maximum-Likelihood Estimation of the Parameters of the Gompertz Survival Function. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *19*(2), 152–159. <https://doi.org/10.2307/2346545>
- Gebelein, H. (1941). Das statistische Problem der Korrelation als Variations- und Eigenwertproblem und sein Zusammenhang mit der Ausgleichsrechnung. *ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik*, *21*(6), 364–379. <https://doi.org/10.1002/zamm.19410210604>
- van der Geest, P. A. G. (1998). An algorithm to generate samples of multi-variate distributions with correlated marginals. *Computational Statistics & Data Analysis*, *27*(3), 271–289. [https://doi.org/10.1016/S0167-9473\(98\)00005-X](https://doi.org/10.1016/S0167-9473(98)00005-X)
- Genton, M. G. (2007). Separable approximations of space-time covariance matrices. *Environmetrics*, *18*(7), 681–695. <https://doi.org/10.1002/env.854>
- Genton, M. G., & Kleiber, W. (2015). Cross-Covariance Functions for Multivariate Geostatistics. *Statistical Science*, *30*(2), 147–163. <https://doi.org/10.1214/14-STS487>
- Gneiting, T., & Schlather, M. (2004). Stochastic Models That Separate Fractal Dimension and the Hurst Effect. *SIAM Review*, *46*(2), 269–282. <https://doi.org/10.1137/S0036144501394387>

- Gneiting, Tilmann. (1999). Correlation functions for atmospheric data analysis. *Quarterly Journal of the Royal Meteorological Society*, 125(559), 2449–2464. <https://doi.org/10.1002/qj.49712555906>
- Gneiting, Tilmann. (2000). Power-law correlations, related models for long-range dependence and their simulation. *Journal of Applied Probability*, 37(4), 1104–1109. <https://doi.org/10.1239/jap/1014843088>
- Gneiting, Tilmann. (2002). Nonseparable, Stationary Covariance Functions for Space–Time Data. *Journal of the American Statistical Association*, 97(458), 590–600. <https://doi.org/10.1198/016214502760047113>
- Gneiting, Tilmann, Genton, M., & Guttorp, P. (2006). Geostatistical Space–Time Models, Stationarity, Separability, and Full Symmetry. In B. Finkenst $\sqrt{\text{dt}}$, L. Held, & V. Isham (Eds.), *Statistical Methods for Spatio-Temporal Systems* (Vol. 107, pp. 151–175). Chapman and Hall/CRC. <https://doi.org/10.1201/9781420011050.ch4>
- Goldstein, R. S. (2000). The Term Structure of Interest Rates as a Random Field. *The Review of Financial Studies*, 13(2), 365–384. <https://doi.org/10.1093/rfs/13.2.365>
- Gompertz, B. (1815). On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies. *Proceedings of the Royal Society of London Series I*, 2, 252–253.
- Gupta, V. K., & Waymire, E. C. (1993). A Statistical Analysis of Mesoscale Rainfall as a Random Cascade. *Journal of Applied Meteorology*, 32(2), 251–267. [https://doi.org/10.1175/1520-0450\(1993\)032<0251:ASAOMR>2.0.CO;2](https://doi.org/10.1175/1520-0450(1993)032<0251:ASAOMR>2.0.CO;2)
- Hardy, H. H., & Beier, R. A. (1994). *Fractals in Reservoir Engineering*. World Scientific.
- He, J., Christakos, G., Wu, J., Jankowski, P., Langousis, A., Wang, Y., et al. (2019). Probabilistic logic analysis of the highly heterogeneous spatiotemporal HFRS incidence distribution in Heilongjiang province (China) during 2005–2013. *PLOS Neglected Tropical Diseases*, 13(1), e0007091. <https://doi.org/10.1371/journal.pntd.0007091>
- Higham, N. J. (2002). Computing the nearest correlation matrix—a problem from finance. *IMA Journal of Numerical Analysis*, 22(3), 329–343. <https://doi.org/10.1093/imanum/22.3.329>
- Hurst, H. E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116, 770–808.
- Jinno, K., Kawamura, A., Berndtsson, R., Larson, M., & Niemczynowicz, J. (1993). Real-time rainfall prediction at small space-time scales using a two-dimensional stochastic advection-diffusion model. *Water Resources Research*, 29(5), 1489–1504. <https://doi.org/10.1029/92WR02849>
- Jona-Lasinio, G., Gelfand, A., & Jona-Lasinio, M. (2012). SPATIAL ANALYSIS OF WAVE DIRECTION DATA USING WRAPPED GAUSSIAN PROCESSES. *The Annals of Applied Statistics*, 6(4), 1478–1498.
- Karimaghloo, Z., Arnold, D. L., & Arbel, T. (2016). Adaptive multi-level conditional random fields for detection and segmentation of small enhanced pathology in medical images. *Medical Image Analysis*, 27, 17–30. <https://doi.org/10.1016/j.media.2015.06.004>
- Kawamura, A., Jinno, K., Berndtsson, R., & Furukawa, T. (1996). Parameterization of rain cell properties using an advection-diffusion model and rain gage data. *Atmospheric Research*, 42(1), 67–73. [https://doi.org/10.1016/0169-8095\(95\)00053-4](https://doi.org/10.1016/0169-8095(95)00053-4)
- Kleiber, W. (2016). High resolution simulation of nonstationary Gaussian random fields. *Computational Statistics & Data Analysis*, 101, 277–288. <https://doi.org/10.1016/j.csda.2016.03.005>
- Kolovos, A., Christakos, G., Hristopulos, D. T., & Serre, M. L. (2004). Methods for generating non-separable spatiotemporal covariance models with potential environmental applications. *Advances in Water Resources*, 27(8), 815–830. <https://doi.org/10.1016/j.advwatres.2004.04.002>
- Kumar, P., & Foufoula-Georgiou, E. (1997). Wavelet analysis for geophysical applications. *Reviews of Geophysics*, 35(4), 385–412. <https://doi.org/10.1029/97RG00427>
- Kyriakidis, P. C., & Journel, A. G. (1999). Geostatistical Space–Time Models: A Review. *Mathematical Geology*, 31(6), 651–684. <https://doi.org/10.1023/A:1007528426688>
- Lancaster, H. O. (1957). Some Properties of the Bivariate Normal Distribution Considered in the Form of a Contingency Table. *Biometrika*, 44(1/2), 289–292. <https://doi.org/10.2307/2333274>
- Lang, A., & Potthoff, J. (2011). Fast simulation of Gaussian random fields. *Monte Carlo Methods and Applications*, 17(3), 195–214. <https://doi.org/10.1515/mcma.2011.009>
- Langousis, A., & Veneziano, D. (2007). Intensity-duration-frequency curves from scaling representations of rainfall. *Water Resources Research*, 43(2). <https://doi.org/10.1029/2006WR005245>
- Langousis, A., Veneziano, D., Furcolo, P., & Lepore, C. (2009). Multifractal rainfall extremes: Theoretical analysis and practical estimation. *Chaos, Solitons & Fractals*, 39(3), 1182–1194. <https://doi.org/10.1016/j.chaos.2007.06.004>
- Lantuejoul, C. (2013). *Geostatistical Simulation: Models and Algorithms*. Springer Science & Business Media.
- Leblois, E., & Creutin, J.-D. (2013). Space-time simulation of intermittent rainfall with prescribed advection field: Adaptation of the turning band method. *Water Resources Research*, 49(6), 3375–3387. <https://doi.org/10.1002/wrcr.20190>
- Lescourret, L., & Robert, C. Y. (2006). Extreme dependence of multivariate catastrophic losses. *Scandinavian Actuarial Journal*, 2006(4), 203–225. <https://doi.org/10.1080/03461230600889645>

- Li, B., Eriksson, M., Srinivasan, R., & Sherman, M. (2008). A geostatistical method for Texas NexRad data calibration. *Environmetrics*, 19(1), 1–19. <https://doi.org/10.1002/env.848>
- Li, S. T., & Hammond, J. L. (1975). Generation of Pseudorandom Numbers with Specified Univariate Distributions and Correlation Coefficients. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-5*(5), 557–561. <https://doi.org/10.1109/TSMC.1975.5408380>
- Liou, J.-J., Su, Y.-F., Chiang, J.-L., & Cheng, K.-S. (2011). Gamma random field simulation by a covariance matrix transformation method. *Stochastic Environmental Research and Risk Assessment*, 25(2), 235–251. <https://doi.org/10.1007/s00477-010-0434-8>
- Lovejoy, S., & Mandelbrot, B. B. (1985). Fractal properties of rain, and a fractal model. *Tellus A: Dynamic Meteorology and Oceanography*, 37(3), 209–232. <https://doi.org/10.3402/tellusa.v37i3.11668>
- Lovejoy, Shaun, & Schertzer, D. (1990). Multifractals, universality classes and satellite and radar measurements of cloud and rain fields. *Journal of Geophysical Research: Atmospheres*, 95(D3), 2021–2034.
- Ma, C. (2002). Spatio-Temporal Covariance Functions Generated by Mixtures. *Mathematical Geology*, 34(8), 965–975. <https://doi.org/10.1023/A:1021368723926>
- Mandelbrot, B. B., & Wallis, J. R. (1968). Noah, Joseph, and Operational Hydrology. *Water Resources Research*, 4(5), 909–918. <https://doi.org/10.1029/WR004i005p00909>
- Mandelbrot, B. B., & Wallis, J. R. (1969). Some long-run properties of geophysical records. *Water Resources Research*, 5(2), 321–340. <https://doi.org/10.1029/WR005i002p00321>
- Mantoglou, A., & Wilson, J. L. (1982). The Turning Bands Method for simulation of random fields using line generation by a spectral method. *Water Resources Research*, 18(5), 1379–1394. <https://doi.org/10.1029/WR018i005p01379>
- Markonis, Y., Moustakis, Y., Nasika, C., Sychova, P., Dimitriadis, P., Hanel, M., et al. (2018). Global estimation of long-term persistence in annual river runoff. *Advances in Water Resources*, 113, 1–12. <https://doi.org/10.1016/j.advwatres.2018.01.003>
- Martin, R. J., & Walker, A. M. (1997). A power-law model and other models for long-range dependence. *Journal of Applied Probability*, 34(3), 657–670. <https://doi.org/10.2307/3215092>
- Matalas, N. C. (1967). Mathematical assessment of synthetic hydrology. *Water Resources Research*, 3(4), 937–945.
- Maung, K. (1941). Measurement of association in a contingency table with special reference to the pigmentation of hair and eye colours of Scottish school children. *Annals of Eugenics*, 11(1), 189–223.
- Mejía, J. M., & Rodríguez-Iturbe, I. (1974). On the synthesis of random field sampling from the spectrum: An application to the generation of hydrologic spatial processes. *Water Resources Research*, 10(4), 705–711. <https://doi.org/10.1029/WR010i004p00705>
- van Montfort, M. A. J. (1990). Sliding maxima. *Journal of Hydrology*, 118(1–4), 77–85. [https://doi.org/10.1016/0022-1694\(90\)90251-R](https://doi.org/10.1016/0022-1694(90)90251-R)
- Nelsen, R. B. (2006). *An Introduction to Copulas* (Second). Springer-Verlag, New York.
- Nerantzaki, S., & Papalexiou, S. M. (2019). Tails of Extremes: Advancing a Graphical Method and Harnessing Big Data to Assess Precipitation Extremes. *Advances in Water Resources*, 103448. <https://doi.org/10.1016/j.advwatres.2019.103448>
- Nevat, I., Peters, G. W., Septier, F., & Matsui, T. (2015). Estimation of Spatially Correlated Random Fields in Heterogeneous Wireless Sensor Networks. *IEEE Transactions on Signal Processing*, 63(10), 2597–2609. <https://doi.org/10.1109/TSP.2015.2412917>
- Papalexiou, S.M., Koutsoyiannis, D., & Makropoulos, C. (2013). How extreme is extreme? An assessment of daily rainfall distribution tails. *Hydrology and Earth System Sciences*, 17(2), 851–862.
- Papalexiou, S.M. (2018). Unified theory for stochastic modelling of hydroclimatic processes: Preserving marginal distributions, correlation structures, and intermittency. *Advances in Water Resources*, 115, 234–252. <https://doi.org/10.1016/j.advwatres.2018.02.013>
- Papalexiou, S.M., & Koutsoyiannis, D. (2012). Entropy based derivation of probability distributions: A case study to daily rainfall. *Advances in Water Resources*, 45, 51–57. <https://doi.org/10.1016/j.advwatres.2011.11.007>
- Papalexiou, S.M., & Koutsoyiannis, D. (2016). A global survey on the seasonal variation of the marginal distribution of daily precipitation. *Advances in Water Resources*, 94, 131–145. <https://doi.org/10.1016/j.advwatres.2016.05.005>
- Papalexiou, S.M., Dialynas, Y. G., & Grimaldi, S. (2016). Hershfield factor revisited: Correcting annual maximum precipitation. *Journal of Hydrology*, 542, 884–895. <https://doi.org/10.1016/j.jhydrol.2016.09.058>
- Papalexiou, S.M., AghaKouchak, A., & Foufoula-Georgiou, E. (2018). A Diagnostic Framework for Understanding Climatology of Tails of Hourly Precipitation Extremes in the United States. *Water Resources Research*. <https://doi.org/10.1029/2018WR022732>
- Papalexiou, S.M., Markonis, Y., Lombardo, F., AghaKouchak, A., & Foufoula-Georgiou, E. (2018). Precise Temporal Disaggregation Preserving Marginals and Correlations (DiPMaC) for Stationary and Nonstationary Processes. *Water Resources Research*. <https://doi.org/10.1029/2018WR022726>

- Paschalis, A., Molnar, P., Fatichi, S., & Burlando, P. (2013). A stochastic model for high-resolution space-time precipitation simulation. *Water Resources Research*, 49(12), 8400–8417. <https://doi.org/10.1002/2013WR014437>
- Pegram, G. G. S., & Clothier, A. N. (2001). High resolution space-time modelling of rainfall: the “String of Beads” model. *Journal of Hydrology*, 241(1), 26–41. [https://doi.org/10.1016/S0022-1694\(00\)00373-5](https://doi.org/10.1016/S0022-1694(00)00373-5)
- Peleg, N., Fatichi, S., Paschalis, A., Molnar, P., & Burlando, P. (2017). An advanced stochastic weather generator for simulating 2-D high-resolution climate variables. *Journal of Advances in Modeling Earth Systems*, 9(3), 1595–1627. <https://doi.org/10.1002/2016MS000854>
- Porcu, E., Bevilacqua, M., & Genton, M. G. (2016). Spatio-Temporal Covariance and Cross-Covariance Functions of the Great Circle Distance on a Sphere. *Journal of the American Statistical Association*, 111(514), 888–898. <https://doi.org/10.1080/01621459.2015.1072541>
- Rényi, A. (1959). On measures of dependence. *Acta Mathematica Academiae Scientiarum Hungarica*, 10(3), 441–451. <https://doi.org/10.1007/BF02024507>
- Robert, C. Y. (2013). Some new classes of stationary max-stable random fields. *Statistics & Probability Letters*, 83(6), 1496–1503. <https://doi.org/10.1016/j.spl.2013.02.017>
- Ronold Knut O. (1990). Random Field Modeling of Foundation Failure Modes. *Journal of Geotechnical Engineering*, 116(4), 554–570. [https://doi.org/10.1061/\(ASCE\)0733-9410\(1990\)116:4\(554\)](https://doi.org/10.1061/(ASCE)0733-9410(1990)116:4(554))
- Salvadori, G., Michele, C. D., Kottegoda, N. T., & Rosso, R. (2007). *Extremes in nature: An approach using copulas*. Dordrecht, The Netherlands: Springer.
- Sampson, P. D., & Guttorp, P. (1992). Nonparametric Estimation of Nonstationary Spatial Covariance Structure. *Journal of the American Statistical Association*, 87(417), 108–119. <https://doi.org/10.2307/2290458>
- Schertzer, D., & Lovejoy, S. (1987). Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes. *Journal of Geophysical Research: Atmospheres*, 92(D8), 9693–9714. <https://doi.org/10.1029/JD092iD08p09693>
- Schlather, M. (1999). An introduction to positive de nite functions and to unconditional simulation of random elds.
- Schlather, Martin. (2002). Models for Stationary Max-Stable Random Fields. *Extremes*, 5(1), 33–44. <https://doi.org/10.1023/A:1020977924878>
- Schlather, Martin. (2012). Construction of Covariance Functions and Unconditional Simulation of Random Fields. In E. Porcu, J. Montero, & M. Schlather (Eds.), *Advances and Challenges in Space-time Modelling of Natural Events* (pp. 25–54). Springer Berlin Heidelberg.
- Serinaldi, F., & Kilsby, C. G. (2014). Rainfall extremes: Toward reconciliation after the battle of distributions. *Water Resources Research*, 50(1), 336–352.
- Serinaldi, F., & Kilsby, C. G. (2016). A Blueprint for Full Collective Flood Risk Estimation: Demonstration for European River Flooding. *Risk Analysis*. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/risa.12747/full>
- Serinaldi, F., Bárdossy, A., & Kilsby, C. G. (2015). Upper tail dependence in rainfall extremes: would we know it if we saw it? *Stochastic Environmental Research and Risk Assessment*, 29(4), 1211–1233. <https://doi.org/10.1007/s00477-014-0946-8>
- Stacy, E. W. (1962). A Generalization of the Gamma Distribution. *The Annals of Mathematical Statistics*, 33(3), 1187–1192.
- Sun, Y., Li, B., & Genton, M. G. (2012). Geostatistics for Large Datasets. In E. Porcu, J. Montero, & M. Schlather (Eds.), *Advances and Challenges in Space-time Modelling of Natural Events* (pp. 55–77). Springer Berlin Heidelberg.
- Tadikamalla, P. R. (1980). A Look at the Burr and Related Distributions. *International Statistical Review / Revue Internationale de Statistique*, 48(3), 337–344.
- Tyralis, H., & Langousis, A. (2019). Estimation of intensity–duration–frequency curves using max-stable processes. *Stochastic Environmental Research and Risk Assessment*, 33(1), 239–252. <https://doi.org/10.1007/s00477-018-1577-2>
- Vanmarcke, E. (1983). Random Fields. *Random Fields*, 372.
- Veneziano, D., Langousis, A., & Furcolo, P. (2006). Multifractality and rainfall extremes: A review. *Water Resources Research*, 42(6). <https://doi.org/10.1029/2005WR004716>
- Wackernagel, H. (2003). *Multivariate Geostatistics: An Introduction with Applications* (3rd ed.). Berlin Heidelberg: Springer-Verlag. Retrieved from <https://www.springer.com/gp/book/9783540441427>
- Winkler, G. (1995). *Image Analysis, Random Fields and Dynamic Monte Carlo Methods: A Mathematical Introduction*. Berlin Heidelberg: Springer-Verlag. <https://doi.org/10.1007/978-3-642-97522-6>
- Wolpert, R. L., & Ickstadt, K. (1998). Poisson/gamma random field models for spatial statistics. *Biometrika*, 85(2), 251–267. <https://doi.org/10.1093/biomet/85.2.251>
- Youngman, B. D., & Stephenson, D. B. (2016). A geostatistical extreme-value framework for fast simulation of natural hazard events. In *Proc. R. Soc. A* (Vol. 472, p. 20150855). The Royal Society. Retrieved from <http://rspa.royalsocietypublishing.org/content/472/2189/20150855.abstract>

Annex A

A.1 Product moments

Here we provide the raw moments of order q for the Generalized Gamma (GG), Beta of the Second Kind (BII), Burr type XII (B_{rXII}), and Burr type III (B_{rIII}) distributions shown in Equations (1) to (4).

$$m_{GG}(q) = \beta^q \Gamma\left(\frac{q + \gamma_1}{\gamma_2}\right) / \Gamma\left(\frac{\gamma_1}{\gamma_2}\right) \quad (A.1)$$

$$m_{BII}(q) = \left(\frac{\beta}{\gamma_2}\right)^q \frac{\Gamma(q + \gamma_1) \Gamma(1/\gamma_2 - q)}{\Gamma(\gamma_1) \Gamma(1/\gamma_2)} \quad (A.2)$$

$$m_{B_{rXII}}(q) = \frac{\beta^q \gamma_2^{-1-q/\gamma_1}}{\gamma_1} B\left(\frac{q}{\gamma_1} + 1, \frac{1 - q\gamma_2}{\gamma_1\gamma_2}\right) \quad (A.3)$$

$$m_{B_{rIII}}(q) = \beta^q \gamma_1^{1-\gamma_2 q} \gamma_2 B(1 - q\gamma_2, (q + \gamma_1) \gamma_2) \quad (A.4)$$

Where $\Gamma(a) = \int_0^\infty t^{a-1} \exp(-t) dt$ is the Gamma function, and $B(a, b) = \int_0^1 t^{a-1} (1-t)^{b-1} dt$ is the Beta function. The mean μ and the variance σ^2 can be estimated by $\mu = m(1)$ and $\sigma^2 = m(2) - \mu^2$.

A.2 Anisotropy

As mentioned in the main text, anisotropy might be needed in the simulation of high-resolution space-time rainfall fields as well as other variables whose properties exhibit preferential directions. A common way to introduce anisotropy is by using a transformed coordinate space $\tilde{\mathbf{s}} = A R \mathbf{s}$, where

$$A = \begin{pmatrix} 1/\phi_1 & 0 \\ 0 & 1/\phi_2 \end{pmatrix} \quad (A.5)$$

$$R = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (A.6)$$

and $\mathbf{s} = (s_i, s_j)^T$. This transformation corresponds to scaling by $1/\phi_1$ and $1/\phi_2$ in two orthogonal directions (e.g. longitude and latitude), and a counter-clockwise rotation through angle θ (e.g. Youngman & Stephenson, 2016).

A.3 Additional Figures

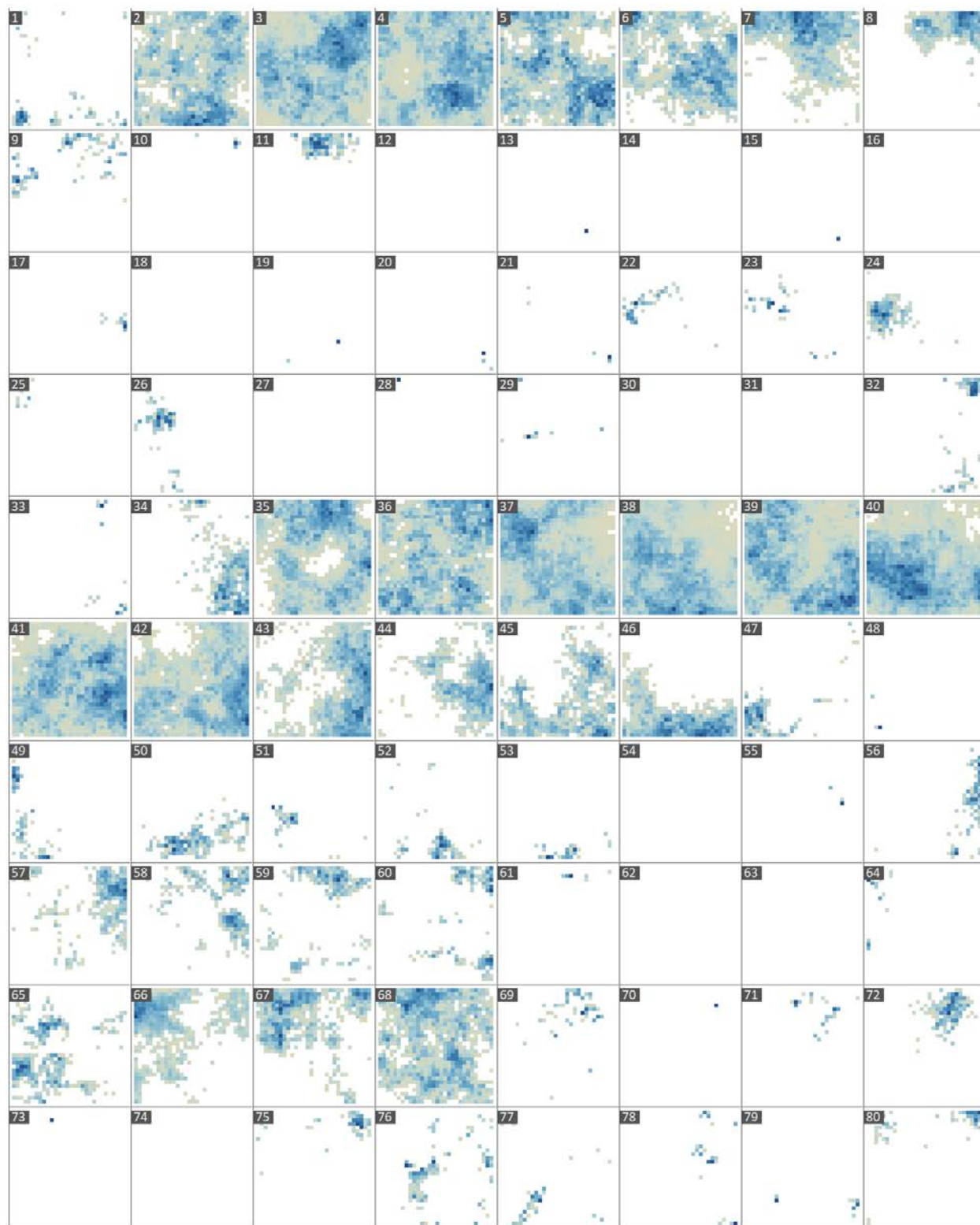


Figure A1. Simulation of 80 precipitation fields (see Section 3.1).

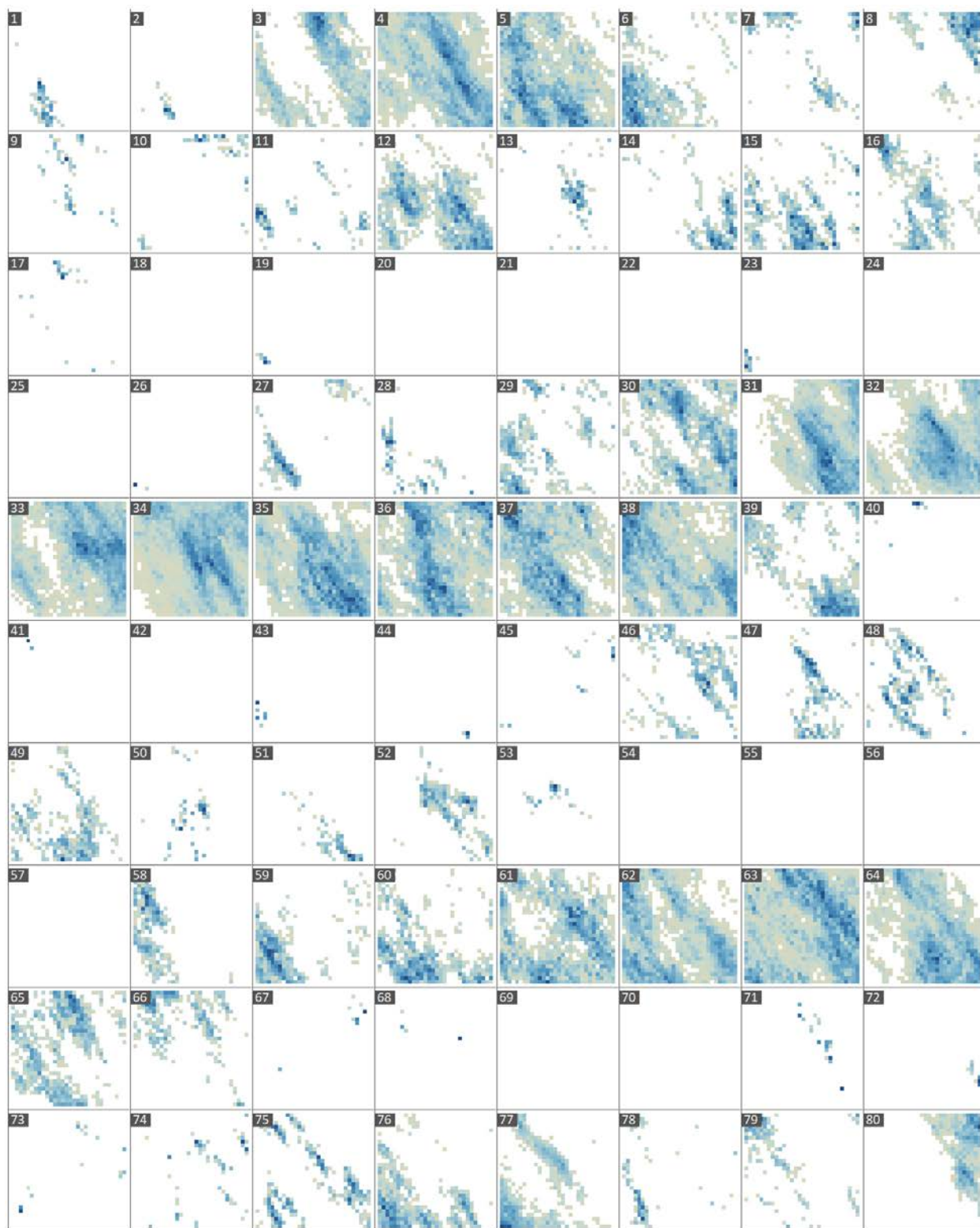


Figure A2. Simulation of 80 precipitation fields based on the same model as in Figure A1 with added anisotropy (anisotropy parameters are $\phi_1 = 0.5$, $\phi_2 = 1.5$, $\theta = \pi/6$).

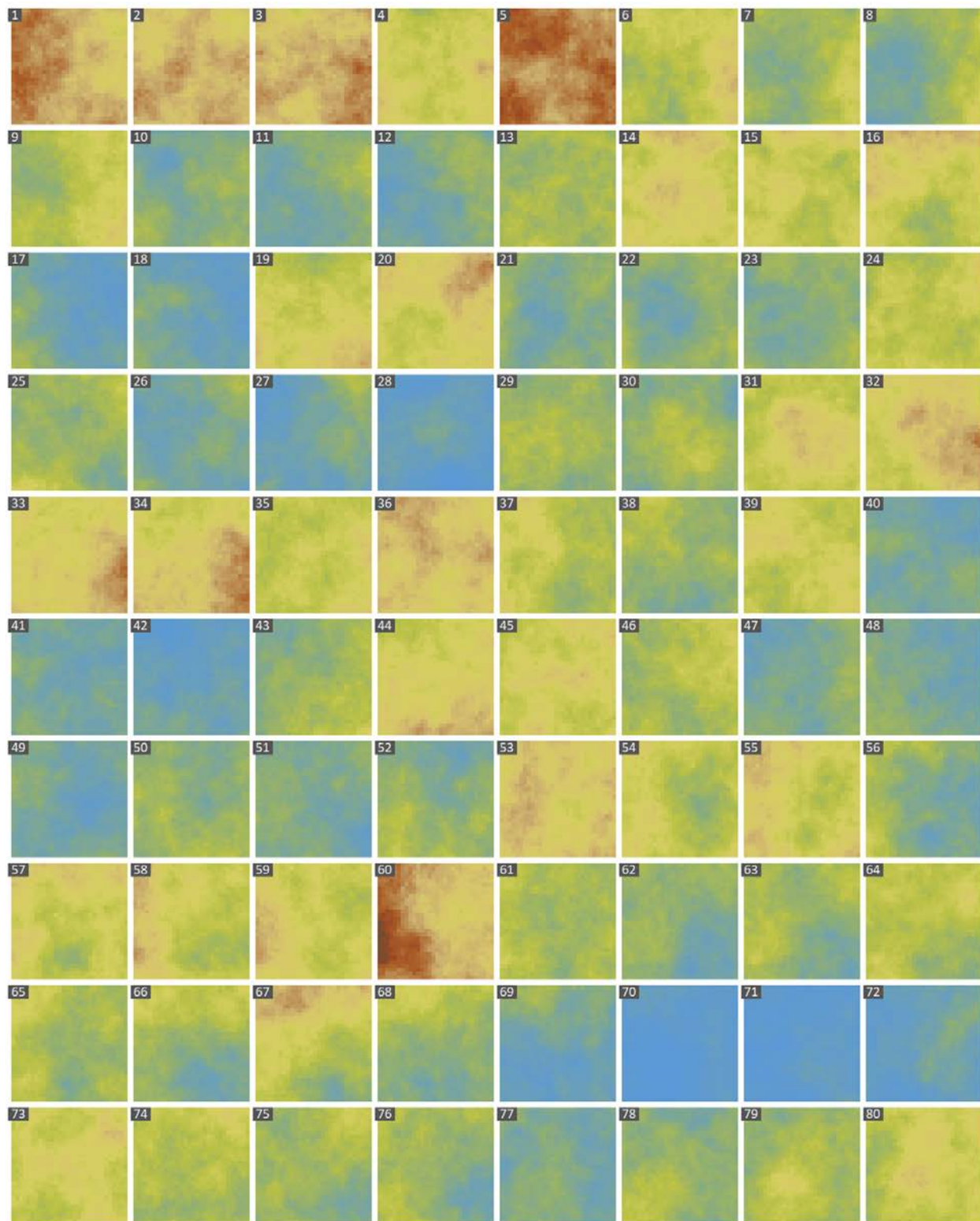


Figure A3. Simulation of 80 relative humidity fields (see Section3.2).

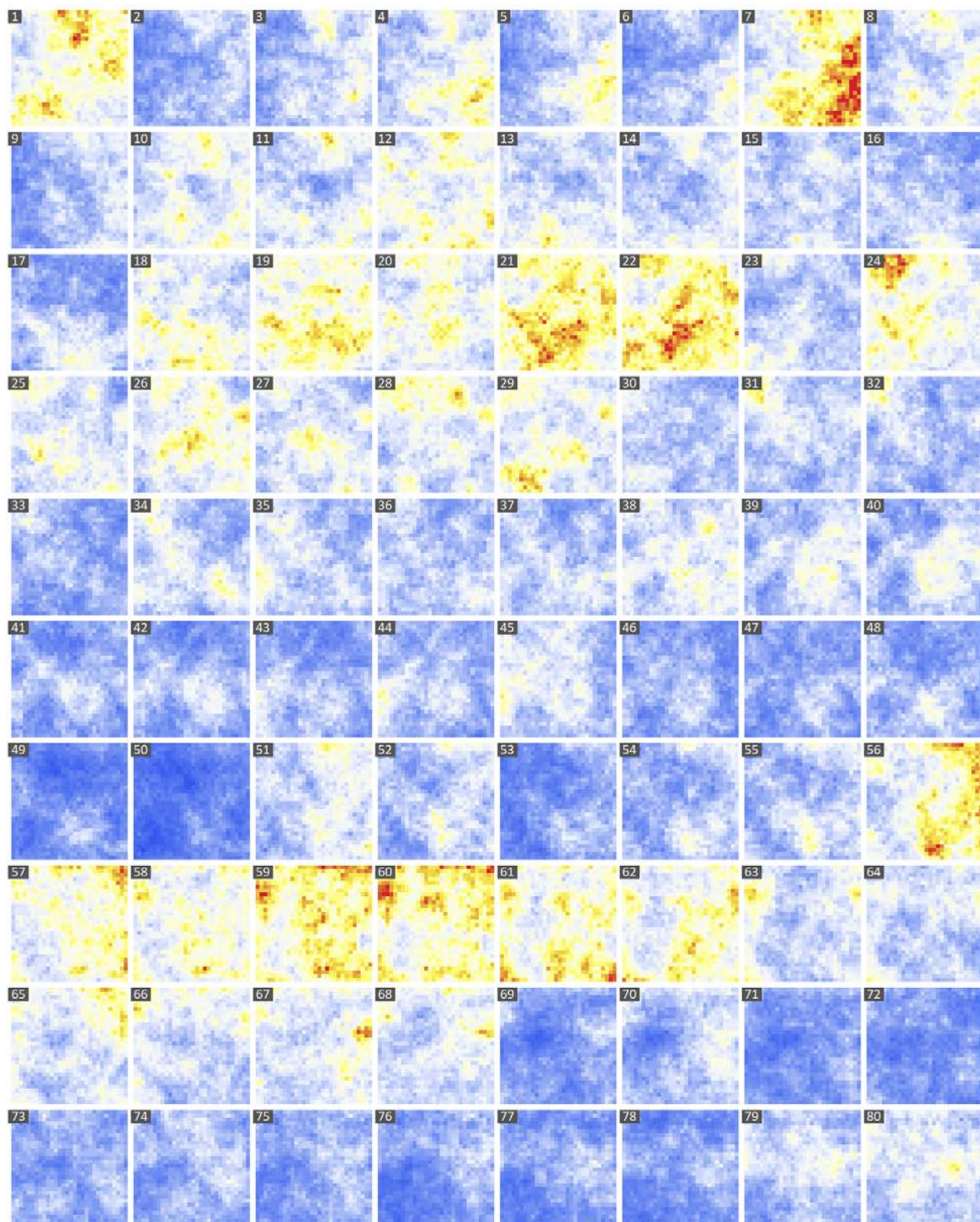


Figure A4. Simulation of 80 daily maximum temperature fields. The strong spatial and temporal correlation is apparent (see Section 3.3).

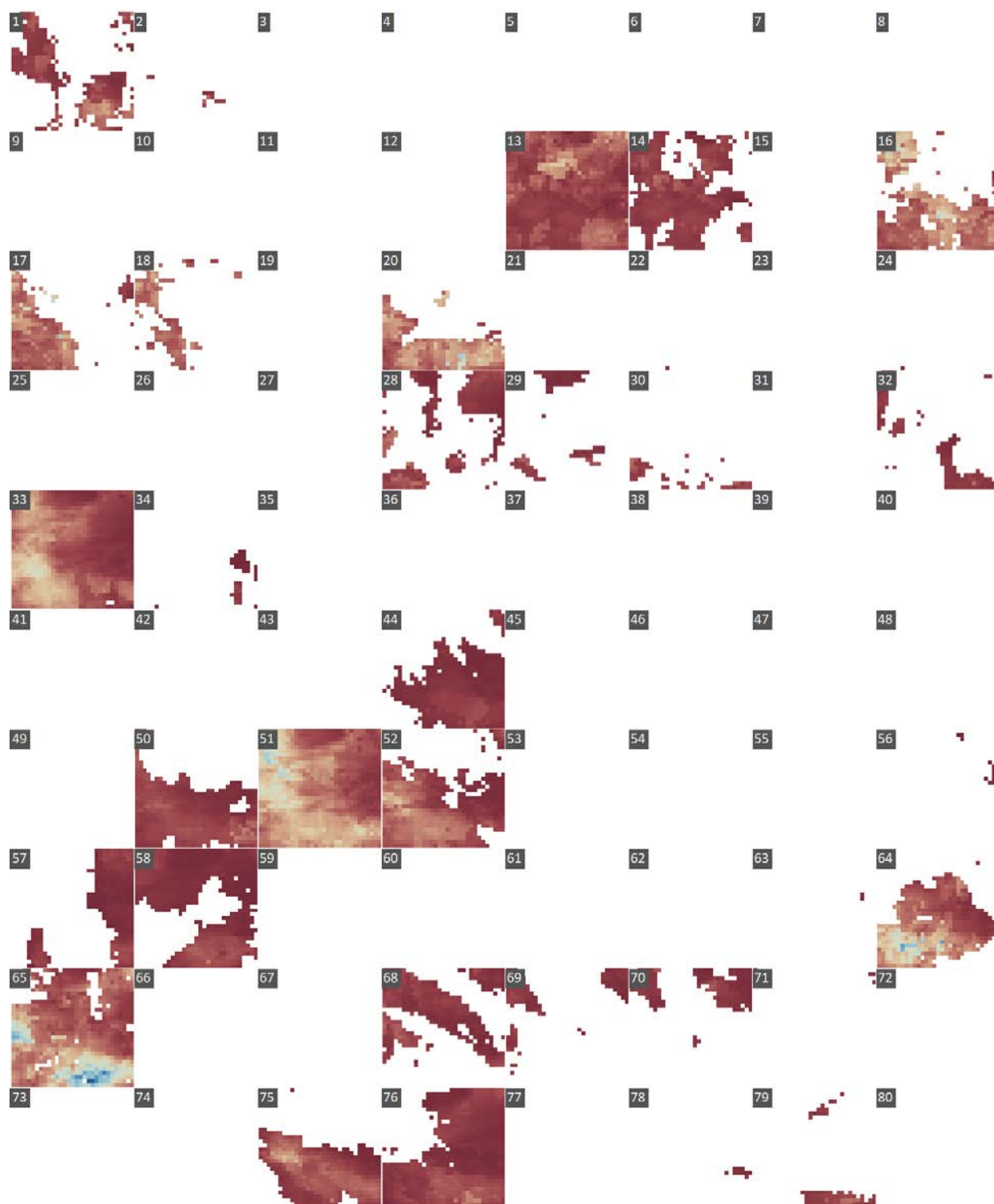


Figure A5. Sequence of 80 observed daily precipitation fields from month January (see Section 4).

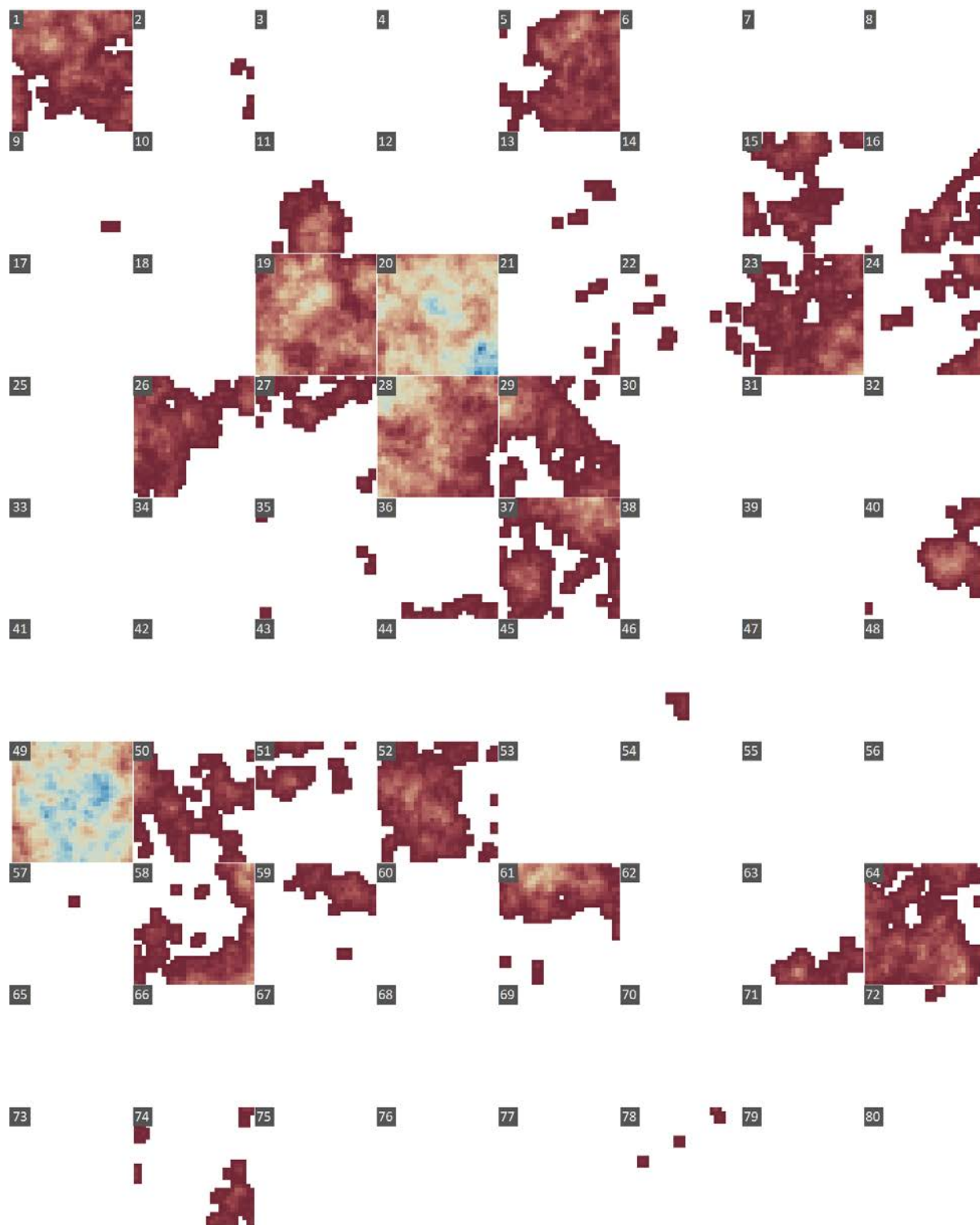


Figure A6. Sequence of 80 simulated daily precipitation fields for month January (see Section 4).